2003

A Genetic Analysis of the Intraspecific Relationships of Tropical Marine Shorefishes Common to Bermuda and the Southeastern Atlantic Coast of the United States

Kelly R. Johnson
College of William and Mary - Virginia Institute of Marine Science

A GENETIC ANALYSIS OF THE INTRASPECIFIC RELATIONSHIPS OF TROPICAL MARINE SHOREFISHES COMMON TO BERMUDA AND THE SOUTHEASTERN ATLANTIC COAST OF THE UNITED STATES

A Thesis

Presented to

The Faculty of the School of Marine Science

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Science

---

by

Kelly R. Johnson

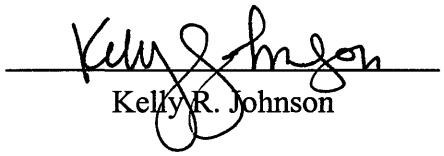
2003

APPROVAL SHEET

This thesis is submitted in partial fulfillment of

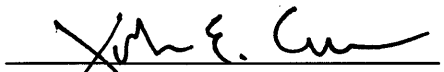
the requirements for the degree of

Master of Science

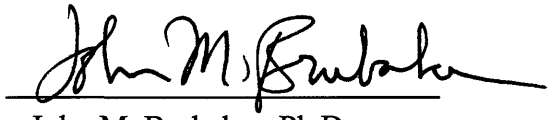


Kelly R. Johnson

Approved, December 2003



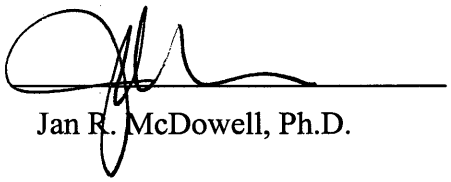
John E. Graves, Ph.D.
Advisor



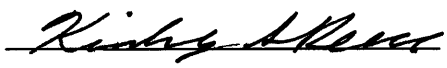
John M. Brubaker, Ph.D.

Bruce B. Collette, Ph.D.

Brian E. Luckhurst, Ph.D.



Jan R. McDowell, Ph.D.



Kimberly S. Reece, Ph.D.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS....iv
LIST OF TABLES....v
LIST OF FIGURES....viii
ABSTRACT....xi
INTRODUCTION....2
Larval Transport....3
Molecular Markers and Phylogeography....9
OBJECTIVES....21
MATERIALS AND METHODS....23
Species Collection....23
DNA Isolation....26
Mitochondrial DNA....27
Nuclear DNA....29
Data Analysis....31
RESULTS....35
Mitochondrial DNA....35
Nuclear DNA....64
DISCUSSION....97
Diversity within Bermuda....97
Diversity between Bermuda and the United States....100
Divergence in the mitochondrial control region....103
Correlation between Markers....106
Divergence Times....107
Anchoa....112
Levels of Taxonomic Divergence....113
CONCLUSIONS....119
LITERATURE CITED....120
VITA....133

Acknowledgments

I would like to acknowledge my major professor, Dr. John Graves, for believing I was up to the challenge of this project and for all of his support and guidance throughout the course of the research. I also wish to thank the members of my Advisory Committee for all of their wonderful assistance, particularly Dr. Jan McDowell for her constant help and friendship, and Dr. Bruce Collette for the knowledge he imparted to me about the world of fishes.

In addition, I would like to thank everyone who assisted in fish collection. In Bermuda: Dr. Bruce Collette (National Marine Fisheries Service, National Museum of Natural History), Dr. Brian Luckhurst (Bermuda Department of Environmental Protection), Dr. Jan McDowell (VIMS), Chrissy Van Hilst (Bermuda Biological Station for Research, Inc.), Jim West, and Kevin Walsh. In the U.S.: Dr. Bruce Collette, Martha Nizinski (National Marine Fisheries Service), Andy Bentley (University of Kansas Natural History Museum), Reade Bonniwell (VIMS Eastern Shore Lab), Jon Lucy (VIMS), Erin McDevitt (Florida Marine Research Intitute), Fred Vose (Florida Marine Research Institute), Michelle Paddock (University of Miami), and Jim Gartland (VIMS). One individual of *Anchoa* spp. was collected by Karl Aiken (The University of the West Indies at Mona, Jamaica) in Jamaica.

LIST OF TABLES

Page

1. Pelagic larval durations and number of individuals collected for seven species of shorefishes common to Bermuda and the United States.....24

2. Length of the control region sequence and number of variable sites for each species.....36

3. The number of 3' end repeats in the control region per sample from each location for the species *Holocentrus adscensionis*.....37

4. Transition/tranversion rations for the entire control region and for an approximately 300 bp section on the 5' tRNA proline end of the control region for each species.....38

5. Number of control region haplotypes, haplotype diversity (h), and mean nucleotide sequence diversity (π) for each species by location.....40

6. Mean nucleotide sequence divergence and net mean nucleotide sequence divergence between Bermudian and U.S. populations of each species for the control region.....42

7. Mean nucleotide sequence divergence and net mean nucleotide sequence divergence between the species of *Anchoa* for the control region.....44

8. Fst and AMOVA results based upon control region sequences for each species.....45

9. Estimated divergence times for each species using published rates of control region evolution.....63

10. Length of the first internal transcribed spacer region and the number of variable sites for each species.....65

11. Number of ITS1 alleles observed per individual in relation to the number of clones sequenced per individual.....66

12. Number of ITS1 alleles, gene diversity, and mean nucleotide sequence diversity (π) for each species by location.....68

13. Mean nucleotide sequence divergence and net mean nucleotide sequence divergence between Bermudian and U.S. populations of each species for the ITS1 region.....70

14. Mean nucleotide sequence divergence and net mean nucleotide sequence divergence between the species of *Anchoa* for the ITS1 region.....71

15. Fst and AMOVA results based upon ITS1 sequences for each species.....73

LIST OF FIGURES

Page

1. Sampling sites of Caribbean reef fish in the southern and northern Caribbean current tracks from Shulman and Bermingham (1995).....5

2. General position of Bermuda in relation to the Gulf Stream 6

3. Map of major surface current flow fields in the southern Middle Atlantic Bight.....8

4. Biogeographic provinces outlined by Briggs (1974) and Ekman (1953) with slight modifications by Muss et al. (2001).....12

5. Diagram of the mitochondrial DNA control region.....15

6. Diagram of one repeat of the multi-copy eukaryotic nuclear ribosomal RNA gene cluster.....18

7. Ranges of three species of western central Atlantic anchovies in the genus *Anchoa*.....25

8. Neighbor-joining tree and minimum spanning network of control region sequences from *Haemulon aurolineatum*.....46

9. Neighbor-joining tree of control region sequences from *Haemulon flavolineatum*.....48

10. Neighbor-joining tree and minimum spanning network of control region sequences from *Haemulon sciurus*.....50

11. Neighbor-joining tree of control region sequences from *Holocentrus adscensionis*.....52

12. Neighbor-joining tree and minimum spanning network of control region sequences from *Holocentrus rufus*.....54

13. Neighbor-joining tree of control region sequences from *Lagodon rhomboides*.....56

14. Neighbor-joining tree of control region sequences from *Lutjanus griseus*.....58

15. Neighbor-joining tree of control region sequences from *Anchoa* species.....60

16. Neighbor-joining tree and minimum spanning network of ITS1 sequences from *Haemulon aurolineatum*.....74

17. Neighbor-joining tree and minimum spanning network of ITS1 sequences from *Haemulon flavolineatum*.....77

18. Neighbor-joining tree and minimum spanning network of ITS1 sequences from *Haemulon sciurus*.....80

19. Neighbor-joining tree of ITS1 sequences from *Holocentrus adscensionis*.....83

20. Neighbor-joining tree and minimum spanning network of ITS1 sequences from *Holocentrus rufus*.....85

21. Neighbor-joining tree of ITS1 sequences from *Lagodon rhomboides*.....88

22. Neighbor-joining tree and minimum spanning network of ITS1 sequences from *Lujtanus griseus*.....90

23. Neighbor-joining tree of ITS1 sequences from *Anchoa* species.....93

Abstract

Bermuda, located in the tropical western Atlantic, includes the northernmost coral reef system in the world. During the last glacial maximum 18,000 years ago, it is believed that low sea levels and cooler water temperatures resulted in the extirpation of Bermudian populations of tropical and subtropical marine shorefishes, including endemic species. Bermuda's present fauna is believed to be the result of re-colonization through larval transport following the retreat of glacial conditions. It has been hypothesized that populations in the southeastern United States served as refugia for species that subsequently recolonized Bermuda through larval transport in the Gulf Stream.

Sequence analysis of the mitochondrial control region and the nuclear ITS1 region was performed to investigate the question of Bermudian phylogeography. Specifically, seven species of common tropical shorefishes were analyzed, eight individuals of each species from Bermuda and eight from the southeastern United States. Haplotype diversities of the control region and gene diversities of ITS1 in Bermudian populations were relatively high, while mean nucleotide sequence diversities of the control region and ITS1 were relatively low. This is indicative of rapid population growth following a period of low effective population size or a founder event. The mitochondrial control region of Bermudian populations exhibits an overall trend of reduced genetic diversity in relation to U.S. populations although the reduction is not significant. *Haemulon flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus* have mitochondrial control region and ITS1 divergence values corresponding to mean divergence times of 2,000 to 21,500 years which may suggest colonization following the last glacial period 18,000 years ago. *Haemulon aurolineatum*,

H. sciurus, and *Holocentrus rufus* have mitochondrial control region and ITS1 divergence values corresponding to mean divergence times of 38,000 to 214,500 years. These species may have colonized Bermuda during a previous interglacial, warming period approximately 120,000-130,000 years ago.

In addition, a preliminary phylogenetic analysis of three species of western central Atlantic anchovy species was performed to investigate the source of Bermudian colonists. The endemic Bermudian anchovy, *Anchoa choerostoma*, is more closely related to a clade containing *A. mitchilli*, with a range restricted to the U.S. coastline, than to *A. hepsetus*, with a range extending throughout the Caribbean. This is consistent with the hypothesis that the origin of Bermudian colonists does not lie in the Caribbean, but may be in populations along the southeastern Atlantic coast of the United States.

A GENETIC ANALYSIS OF THE INTRASPECIFIC RELATIONSHIPS OF TROPICAL MARINE SHOREFISHES COMMON TO BERMUDA AND THE SOUTHEASTERN ATLANTIC COAST OF THE UNITED STATES

Introduction

Situated in the northwestern Sargasso Sea, Bermuda includes the northernmost coral reef system in the world (Smith-Vaniz et al. 1999). This system is characterized by a low level of endemism and a fish fauna that is depauperate relative to other tropical western Atlantic islands (Smith-Vaniz et al. 1999). The entire Bermuda Platform has been affected during times of low glacial sea levels throughout geologic history (Smith-Vaniz et al. 1999). Specifically, during the last glacial maximum 18,000 years before present (Wisconsin glaciation), sea level in the Atlantic Ocean was estimated to have dropped approximately 120-130 m below present levels (Milliman and Emery 1968; Kennett 1982). At that time, the Bermudian reef platform would have been dry with marine life restricted to a narrow band of cooler water surrounding the seamount (Smith-Vaniz et al. 1999). It is believed that both the Gulf Stream and its extension, the North Atlantic Current System, were situated farther to the south during the most recent glacial maximum, resulting in decreased water temperatures near Bermuda (Keffer et al. 1988). The combination of low sea level and decreased sea temperature is believed to have resulted in the extirpation of Bermudian populations of tropical and subtropical marine shorefishes, including endemic species.

Bermuda's present fauna is assumed to be the result of recolonization following the retreat of glacial conditions during the Holocene transgression 15,000 to 17,000 years ago (Kennett 1982; Smith-Vaniz et al. 1999). It has been hypothesized that tropical shorefish populations in the southeastern United States that were not severely impacted by glacial climatic change served as sources for subsequent recolonization of Bermuda through larval transport in the Gulf Stream. The goal of my study is to test this

hypothesis through an analysis of mitochondrial and nuclear DNA sequences from seven tropical shorefish species common to both Bermuda and the southeastern United States.

Larval Transport

Many recent studies have demonstrated that most reef populations are responsible for providing their own recruits (Brogan 1994; Jones 1999; Swearer 1999) and that long-distance transport is generally considered insufficient to replenish and sustain established populations around isolated islands and reefs (Cowen et al. 2000). However, dispersal of larvae in major current systems is known to occur and can facilitate colonization of new areas (Scheltema 1968, 1995; McBride and Able 1998). Schultz and Cowen (1994) suggested that long-distance transport between the Caribbean, the United States, and Bermuda is possible through the Gulf Stream. Specifically, they found that while most simulated transport times between Cape Hatteras, North Carolina and Bermuda took longer than the average pelagic larval durations (PLDs) of Bermudian labroids, on rare occasions, transport times and PLDs were compatible with successful recruitment to Bermuda. In addition, drifter tracks allowed an estimation of water transport time between Cape Hatteras and Bermuda in the range of 33 to 97 days. The authors suggest that rare recruitment events have occurred periodically over evolutionary time scales and may be responsible for the colonization of Bermudaian shorefishes from populations in the United States. A likely route of this transport would begin with the North Equatorial Current.

The North Equatorial Current flows through the islands of the Caribbean in a predominantly southeast to northwest direction (Shulman and Bermingham 1995; Brown

et al. 1998) (Figure 1). Much of this water continues either through the Yucatan Channel or directly into the Straits of Florida to form the Florida Current and subsequently the Gulf Stream (Brown et al. 1998). The Gulf Stream enters the northern Sargasso Sea where cold core rings and eddies occasionally break off and continue on to Bermuda (Schultz and Cowen 1994) (Figure 2). Larvae that are entrained in these currents would thus be potential colonists to Bermudian populations.

Several different spawning strategies exist among tropical island and coral reef fishes that may disperse eggs and larvae into offshore currents, including the species involved in this study. Larger species within many reef fish families tend to migrate offshore to the seaward margin of the reef system where they form large spawning aggregations. For example, the gray snapper, *Lutjanus griseus*, has an extensive spawning migration to the outer reef around the full moon (Thresher 1984; Bortone and Williams 1986). Another example includes species within the family Haemulidae that form offshore spawning migrations at dusk (Thresher 1984). The pinfish, *Lagodon rhomboides*, is also known to migrate offshore in aggregations for spawning (Muncy 1984). Fishes employing these spawning methods tend to have pelagic eggs and larvae that are often entrained in reef edge currents primarily moving offshore (Johannes 1978). Another common form of reproduction, found in smaller reef species, involves the release of pelagic eggs in the surface waters above the reef itself (Johannes 1978). These eggs subsequently become entrained in surface currents. For example, some members of the families Scaridae, Labridae, Mullidae, and Acanthuridae spawn by making a rapid upward dash from the reef with a release of gametes in a fraction of a second at the surface, followed by an equally rapid descent back to the reef (Johannes 1978). A third

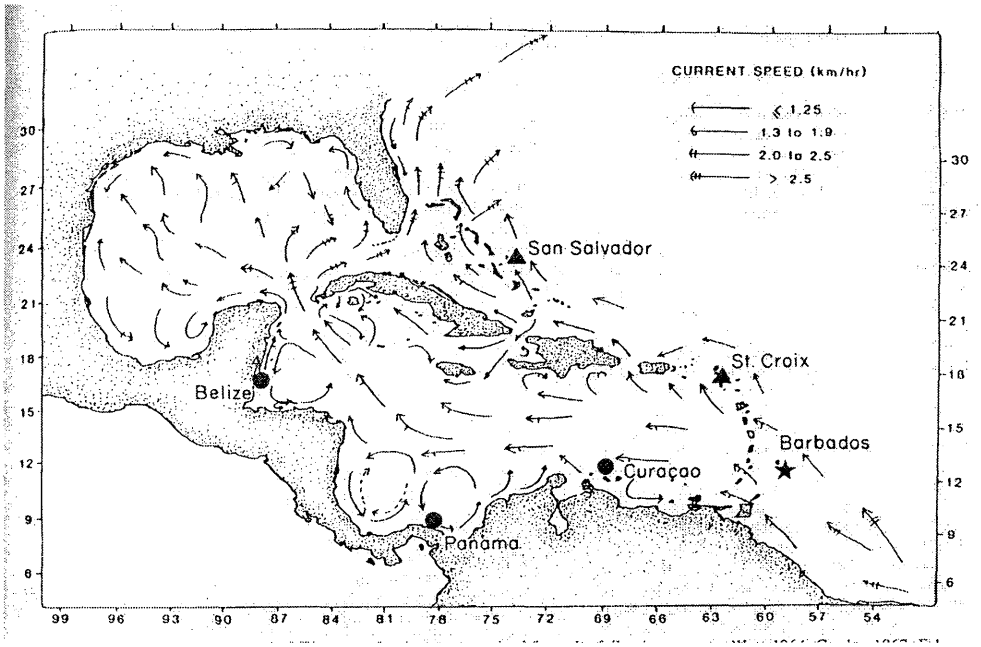


Figure 1. Sampling sites of Caribbean reef fish in the southern and northern Caribbean current tracks from Shulman and Bermingham (1995). Collection sites are as indicated: triangles, northern track; circles, southern track; star, Barbados.

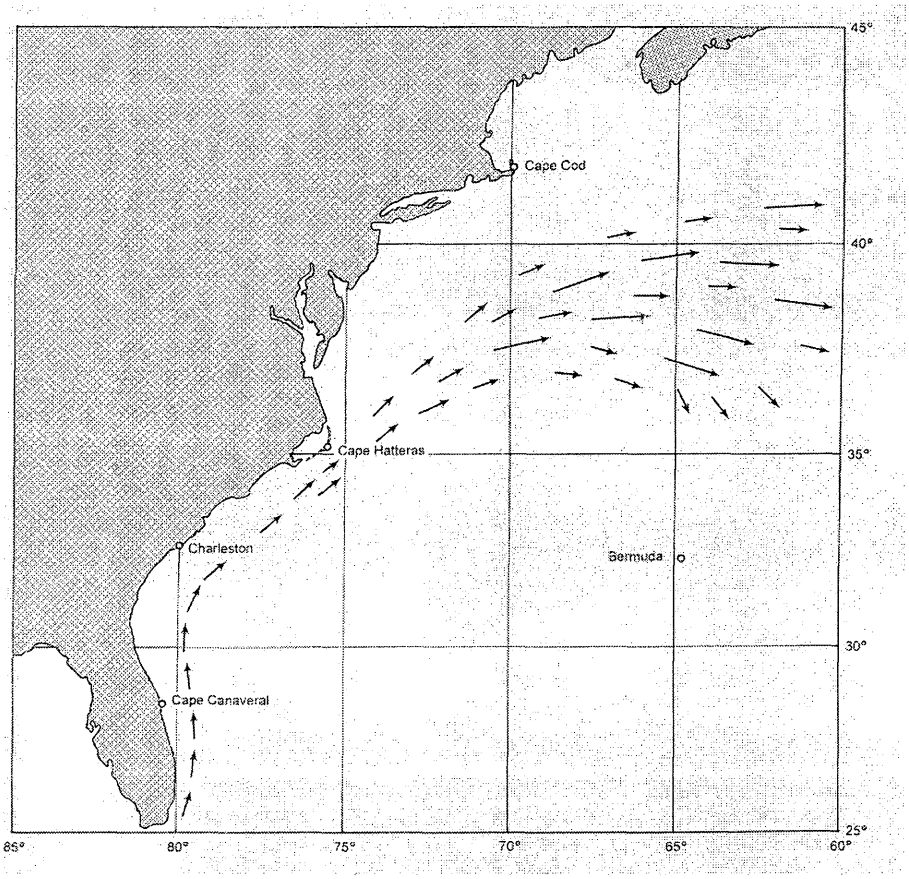


Figure 2. General position of Bermuda in relation to the Gulf Stream (from Smith-Vaniz et al. 1999).

method of spawning, used by many species in the family Sparidae, involves the release of pelagic eggs from the bottom (Thresher 1984). Spawning at night, often at the turn of high tide, as well as spawning on the new or full moon and on the ebbing spring tide, increases the likelihood of an egg reaching offshore gyres near the reef in which development occurs, followed by resettlement of the reef (Johannes 1978). However, in populations along the southeastern United States coast, the offshore currents in which eggs and larvae are initially entrained sometimes are themselves entrained in the major Florida Current eventually forming the Gulf Stream.

In some cases, eggs and larvae of tropical marine shorefishes released along the Atlantic coast of the southeast U.S. must move eastward across the continental shelf before entering the Gulf Stream. Epifanio and Garvine (2001) report that cross-shelf transport in both the Mid-Atlantic and South Atlantic Bights is dominated by two principal forcing agents - the wind and buoyancy differences. Buoyancy driven coastal currents or plumes result when fresh or brackish waters are discharged from rivers and estuaries along the coast. The plumes turn sharply southward and net movement is downshelf, hugging the coastline (Epifanio and Garvine 2001) (Figure 3). Eggs and larvae are commonly advected in plumes moving offshore to the mid-shelf and outer-shelf currents, potentially ending up in the Gulf Stream (Epifanio and Garvine 2001).

Once in the Gulf Stream, larvae can be advected to the northern Sargasso Sea. Mescoscale structures such as cold core rings and eddies are known to pass from the Gulf Stream into the Sargasso Sea near Bermuda (Schultz and Cowen 1994). Larvae and juveniles that have been transported to this point exhibit many behaviors that promote settlement on Bermudian reefs at the end of their pelagic stage. Many late-stage larvae

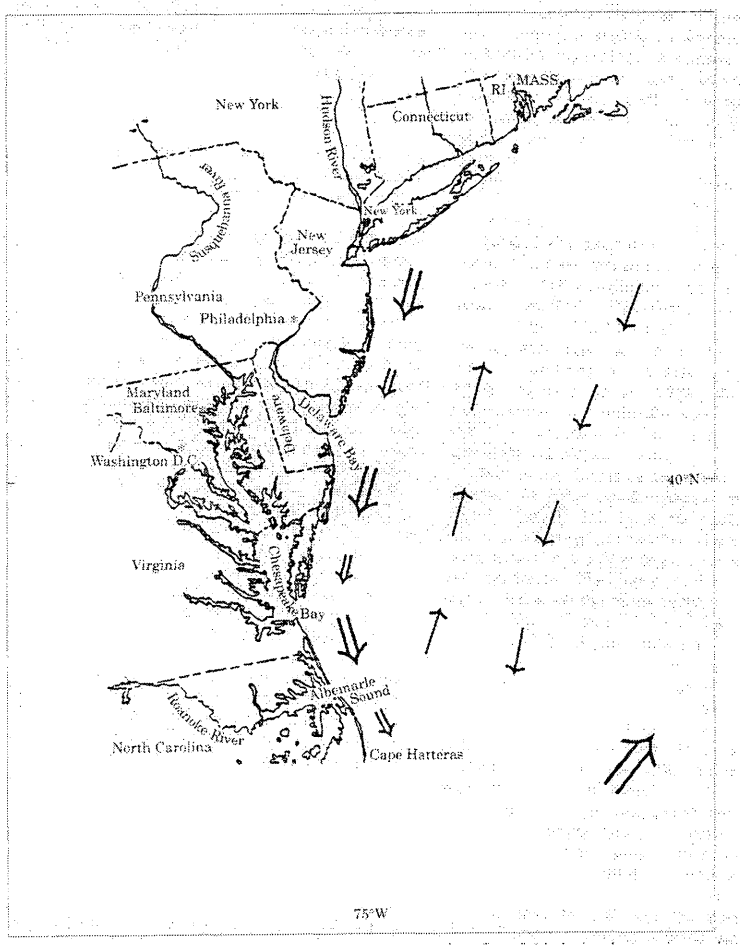


Figure 3. Map of major surface current flow fields in the southern Middle Atlantic Bight. Southward arrows along the coast represent buoyancy-driven flow originating in Hudson, Delaware, and Chesapeake systems. Northward arrows on the mid-shelf represent wind-driven flow associated with upwelling circulation. Southward arrows on outer shelf represent buoyancy-driven flow originating in the Arctic. North-eastward arrow off Cape Hatteras represents the Gulf Stream (from Epifanio and Garvine 2001).

and juveniles settle on the reef during flood tide (Leis 1991). On the windward side of a reef or island, wind-driven surface waters moving towards shore can bring larvae to the reef; on the leeward side, upwelling of deeper, offshore water can bring in more distant larvae (Leis 1991). Ekman transport and density driven flow, important in cross-shelf transport, are also important as mechanisms to promote larval settlement on the reef (Leis 1991).

Molecular Markers and Phylogeography

Hypotheses involving larval transport scenarios such as those described above are commonly evaluated using genetic tools within the discipline of phylogeography. Phylogeography examines how historical processes, such as the breakup of landmasses and changes in oceanic circulation, affect the geographic distribution of genealogical lineages (Avise 2000). Physical processes such as these have effects on populations through factors such as the creation of geographic or ecologic barriers to movement and by affecting routes of larval dispersal and gene flow (Bowen and Avise 1990). These effects can be seen in the geographical distribution of genetic variation within impacted populations. Often, patterns are repeated across several taxa strengthening the phylogeographic signal.

Populations with differing life histories and habitat types are often affected by physical processes in different ways. Freshwater fishes are generally confined to isolated bodies of water formed through historical cycles of drainage and coalescence. Evidence for the phylogeographic history of these fishes is found in strong geographic population structure (Bowen and Avise 1990). At the other extreme, many marine fishes exhibit

life-history patterns involving pelagic stages that are free to disperse in oceanic currents (Bowen and Avise 1990). This dispersal in current systems affects the distribution of genetic variation (Hedgecock 1986; Shulman and Bermingham 1995) and promotes gene flow resulting in much weaker population structure (Palumbi 1994). As current systems change over geological history, the patterns of larval dispersal also presumably change. Several recent phylogeographic studies discuss the importance of larval dispersal on both global and regional scales (eg. Shulman and Bermingham 1995; Bowen et al. 2001; Muss et al. 2001).

The history of the circumtropical marine trumpetfishes provides an example of the effect global changes in oceanic circulation can have on larval dispersal. Bowen et al. (2001) examined three species within the genus *Aulostomus*: *A. chinensis* in the Indian and Pacific Oceans, *A. strigosus* in the eastern Atlantic, and *A. maculatus* in the western Atlantic. The eastern Atlantic species *A. strigosus* and the Indo-Pacific species *A. chinensis* are morphologically indistinguishable, while the western Atlantic species *A. maculatus* is distinguished from the others by a number of meristic characters (Bowen et al. 2001). Bowen et al. (2001) found a deep divergence in the mitochondrial cytochrome *b* gene between *A. maculatus* and *A. chinensis* across the Isthmus of Panama dating to its closure approximately 3-4 million years ago. Populations of *A. chinensis* across the Pacific Ocean share haplotypes and population separation is not significant. The closest relative of *A. chinensis* is *A. strigosus* in the eastern Atlantic, separated by a genetic distance corresponding to approximately 2.5 million years. The separation is due to the "vicariant sundering" of a warm-water connection between the Indian and South Atlantic Oceans around the Cape of Good Hope during the onset of Pliocene glaciation. The

presence of a haplotype unique to *A. strigosus* in coastal Brazil suggests that this species has colonized the southwestern Atlantic possibly leading to the formation of a ring species.

Oceanic barriers and circulation patterns also have detectable effects on population genetic structuring and genetic mixing within individual ocean basins. Muss et al. (2001) investigated the evolution of several species of *Ophioblennius* within five biogeographic provinces in the Atlantic. No mtDNA haplotypes were shared between the provinces represented by ten Atlantic sampling locations, although some were shared between locations within each province (Figure 4). For example, the two Caribbean locations shared the most common haplotype in each sample, but this haplotype was not found in any other sampling location. The mean nucleotide sequence divergence between biogeographic provinces in the Atlantic, the average number of differences per site between sequences in two separate populations, ranged from 0.052 to 0.127. According to Colborn et al. (2001), divergence values greater than 0.05, which is considered to be a minimum value for the designation of separate sister species, suggests deep population structure or deep lineages between these Atlantic provinces. Haplotype diversity, the probability of sampling different haplotypes when two individuals are drawn from the same population, was high (> 0.80) in all populations except the Azores, while nucleotide sequence diversity, the average number of differences per site between sequences within the same population, was low (< 0.013); the Azores had both a low haplotype diversity (0.417) and nucleotide sequence diversity (0.0011). The low levels of both of these measures of variation were suggested to be a vestige of colonization

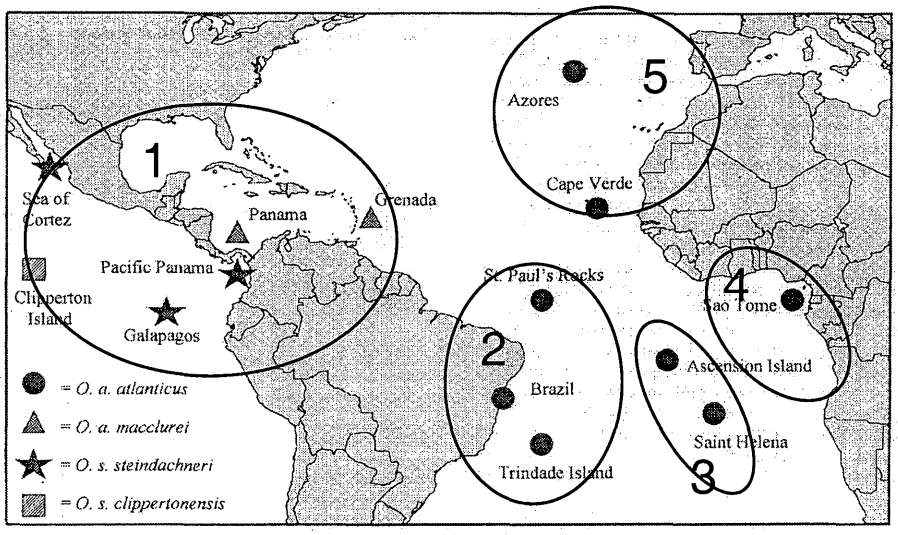


Figure 4. Sample locations for *Ophioblennius*. The biogeographic provinces outlined by Briggs (1974) and Ekman (1953) with slight modifications by Muss et al. (2001) are as follows: 1) tropical east Pacific/tropical west Atlantic or Caribbean; 2) western South Atlantic or Brazil; 3) mid-Atlantic ridge islands; 4) West Africa; 5) Lusitanian province including the Mediterranean (modified from Muss et al. 2001).

following the most recent glacial episode when it is believed that warm water species were extirpated.

Effects of changes in larval dispersal patterns are more difficult to detect on scales smaller than ocean basins. Shulman and Bermingham (1995) studied larval dispersal and gene flow in eight species of Caribbean reef fishes. Six sites were sampled in both the northern and southern Caribbean current tracks (Figure 1). Restriction fragment length polymorphism (RFLP) analysis of the mitochondrial DNA (mtDNA) cytochrome *b* gene revealed no apparent population or phylogeographic structure within species. The most common haplotype was widespread throughout all sample locations in six of the eight species. The remaining haplotypes were closely related with divergence values less than 0.7%. The level of divergence between samples from the southern and northern current tracks was not significantly different than the level found within each current track. The authors suggest gene flow may no longer be present in a widespread manner due to the recent separation of the northern and southern current tracks, but not enough time has elapsed for significant genetic divergence to accumulate.

Phylogeographic hypotheses, such as those described above, have historically been analyzed with a variety of molecular markers including allozymes, mitochondrial DNA (mtDNA), and nuclear DNA. Allozymes were used extensively from the 1970s to the 1990s to examine phylogeography and population structure in a wide variety of fish species, including darters (Turner et al. 1996), sculpins (Riffel and Schreiber 1995) and mesopelagic pearlsides and lanternfish (Baduge Suneetha 2000). However, the greater level of variation found within mitochondrial DNA (mtDNA) has allowed it to become the most widely used method of phylogeographic analysis (Avise 2000).

Mitochondrial DNA

MtDNA is a circular, double-stranded molecule comprising approximately 15,000 to 18,000 nucleotide base pairs and 37 genes in vertebrates (Avise 1994). It is maternally inherited and is haploid, thus in many species it has an effective population size about 1/4 that of nuclear DNA (Taberlet 1996; Avise 2000). The entire mtDNA molecule is considered a single genetic unit (locus) with multiple haplotypes because there is no recombination and all character states are linked due to asexual transmission (Avise 2000). The rate of mtDNA sequence evolution is five to ten times higher than for single-copy nuclear DNA (scnDNA) (Taberlet 1996; Avise 2000). This is hypothesized to result from greater damage caused by a higher proportion of oxygen radicals in the mitochondria (Martin et al. 1992), and from a replication enzyme that lacks self-editing abilities and is prone to a higher rate of nucleotide misincorporation (Brown 1983). In addition, the lower effective population size of mtDNA leads to more rapid fixation of mutations relative to nuclear DNA (Taberlet 1996; Avise 2000).

While the overall rate of sequence evolution is more rapid in mtDNA than nuclear DNA, different regions within both genomes are known to evolve at widely varying rates (Avise 2000). For example, the mitochondrial cytochrome *b* gene, a region encoding the cytochrome b protein involved in the electron transport chain, has several highly conserved sections important in protein function (Palumbi 1996). On the other hand, the mitochondrial control region (Figure 5) contains sections that evolve at a pace four to five times faster than the entire mtDNA molecule (Taberlet 1996). The rapid evolutionary rate of the control region makes it well suited for phylogeographic analyses concerned with timescales of thousands to tens of thousands of years such as those that may be

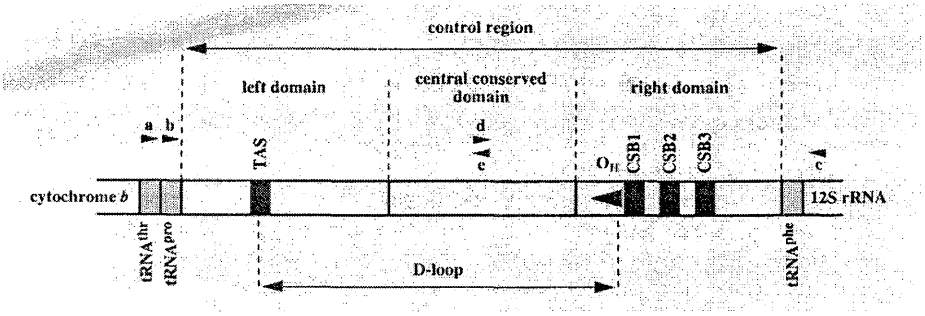


Figure 5. Diagram of the mitochondrial DNA control region (from Taberlet 1996).

encountered between Bermudian and southeastern U.S. marine shorefish populations (Avise 2000).

McMillan and Palumbi (1997) describe variation of the mtDNA control region in considerable detail. In many fish species, including chaetodontids, the control region exhibits a strong transition bias (A ↔ G; C ↔ T); the number of transitions quickly reaches a plateau relative to transversions and subsequent evolution can be obscured by back mutations or multiple mutations at the same site. If an A was replaced with a G and eventually was replaced again with an A, the original change to a G would be obscured. The authors suggest the "striking speed" with which this plateau of transitions is reached means the usefulness of the control region for divergence analyses is restricted to very short periods of geologic time.

Nuclear DNA

Phylogeographic and phylogenetic analyses often use two or more unlinked markers to independently assess evolutionary patterns through an evaluation of the concordance among loci (Schizas et al. 1999). The number of nuclear genes in animal cells is enormous and there are potentially limitless loci to choose for study (Dowling et al. 1996). Nuclear markers frequently used in genetic analyses include both single-copy and multiple copy coding regions, non-transcribed spacer regions, transcribed spacer regions, introns, satellite regions, and randomly amplified polymorphic DNA regions (RAPDs). As in mitochondrial DNA, coding regions are usually quite conserved, most likely due to selective constraints, and may not contain sufficient variation for phylogeographic studies involving short geographic time scales. On the other hand,

intron and transcribed spacer regions are non-coding sequences of DNA that are believed to be have high evolutionary rates due to lack of selective constraints, and thus are sufficiently variable for analyses concerning timescales of thousands to tens of thousands of years (Palumbi 1996).

The internal transcribed spacer (ITS) region of the multi-copy ribosomal RNA (rRNA) gene complex is one such region (Palumbi 1996). The ITS region lies between the small (18S) and large (28S) subunit rRNA genes; it comprises two internal transcribed (but not translated) regions of no known function, ITS1 and ITS2, separated by the 5.8S rRNA gene (Figure 6). The number of copies or paralogues of these ribosomal complexes found in the genome varies among organisms. In humans, for example, hundreds of copies have been found on five pairs of chromosomes (Arnheim 1983). Although variation between the paralogues within an individual has been shown to exist in organisms as diverse as corn (Buckler et al. 1997) and freshwater crayfish (Harris and Crandall 2000), most intragenomic rDNA diversity is quite low due to the process of concerted evolution (Brown et al. 1972). Through the mechanisms of "molecular drive" such as unequal crossing over and gene conversion, individual repeats within the genome evolve in concert, thus resulting in a high degree of similarity, even when found on non-homologous chromosomes (Arnheim 1983; Hillis and Dixon 1991; Elder, Jr. and Turner 1995; Powers et al. 1997). This feature, coupled with the relatively high rate of sequence evolution, make ITS well suited for phylogeographic analyses involving short periods of geologic time. ITS1, in particular, is considered to be more phylogeographically informative as it has been shown to exhibit more variation than ITS2 (Mayer and Soltis 1999; Presa et al. 2002).

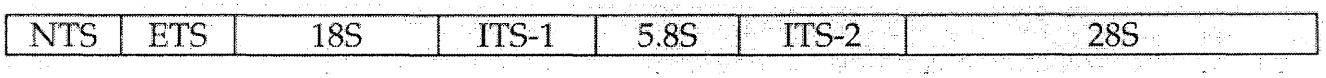


Figure 6. Diagram of one repeat of the multi-copy eukaryotic nuclear ribosomal RNA gene cluster. The non-transcribed spacer region (NTS) separates each repeat unit which includes an external transcribed spacer region (ETS), the 18S small ribosomal subunit gene, the first internal transcribed spacer region (ITS-1), the 5.8S ribosomal subunit gene, the second internal transcribed spacer region (ITS-2), and the 28S large ribosomal subunit gene (from Palumbi 1996).

Molecular Clock

Levels of genetic divergence can be correlated with the time of isolation of populations using the molecular clock theory. This theory holds that the rate of molecular change is constant such that the degree of molecular divergence between two species is linearly related to the time that the two species have been separated (Lessios 1979; Vawter et al. 1980; Hillis et al. 1996). It is rather controversial in that many studies have found heterogeneity of evolutionary rates among various organisms for proteins (Lessios 1979; Bermingham and Lessios 1993), mtDNA (Lessios 1979; Bermingham and Lessios 1993; Bermingham et al. 1997; McMillan and Palumbi 1997; Avise 2000), and nuclear DNA (Li 1993; Avise 1994; Hillis et al. 1996). This has lead to more accepted arguments of the "local" molecular clock and the "episodic" clock. The "local" clock works such that time predictions can be made if the rate of evolution is calibrated for each gene. Among closely related groups with similar life histories and generation times, the rate of evolution for any given gene is likely to be stable (Li 1993; Hillis 1996). An "episodic" clock is one that acknowledges rate variation over short periods of time with rough constancy over long periods of time (Bermingham and Lessios 1993).

One of the most common methods of testing the molecular clock has been to analyze the relative amount of divergence within two or more taxa across the Isthmus of Panama which is known to have closed 3.1 - 3.5 million years ago (Lessios 1979; Vawter et al. 1980). Many studies support the idea of the molecular clock across the Isthmus (Vawter et al. 1980; Knowlton et al. 1993; Knowlton and Weight 1998) and have shown that levels of divergence may even be sensitive enough to detect staggered isolation in

which divergence times for various geminate species pairs across the Isthmus are not simultaneous, but instead reflect different stages in the closure (Knowlton et al. 1993).

In the spirit of the "local" molecular clock, the generally accepted rate of sequence evolution for mtDNA as a whole is approximately 2% sequence variation per million years (Bermingham and Lessios 1993). Estimated rates of sequence evolution for the mitochondrial control region have ranged from 0.93% per million years (Montoya-Burgos 2000) to 108% per million years (McMillan and Palumbi 1997), but are generally accepted to be between 11.5% and 20% per million years (Vigilant et al. 1991; Brown et al. 1993). In a phylogeographic study of brown trout, *Salmo trutta*, using sequences from both the control region and ITS1, Presa et al. (2002) found the rate of sequence evolution of ITS1 to be approximately half that of the control region. The rapid rates of sequence evolution in both the mitochondrial control region and ITS1 make them well-suited to studies of short geological timescales such as the potential extirpation and recolonization of Bermuda believed to have occurred less than 18,000 years ago.

Objectives

The overall goal of this study was to investigate the hypothesis that populations from glacial refugia along the southeastern Atlantic coast of the United States could have served as a source for recolonization of extirpated Bermudian warm-water shorefish populations following the last glacial maximum 18,000 years ago (Wisconsin glaciation). This was accomplished through sequence analysis of both the mitochondrial control region and the first nuclear internal transcribed spacer region (ITS1) from a suite of seven shorefish species common to both Bermuda and the southern United States. These molecular markers were chosen because of their relatively high rates of sequence evolution as the geological timescale in question involves only tens of thousands of years.

The first specific objective of my study was to determine if this suite of species shared a similar level of genetic divergence between Bermudian and U.S. populations which would be consistent with a single recolonization event. It was hypothesized that all species analyzed would exhibit a level of divergence consistent with a separation time of less than 18,000 years.

The second specific objective of my thesis was to determine if the ratio between the rates of sequence evolution of the nuclear ITS1 region and the mitochondrial control region, two independent genetic markers, is constant across all species examined. It was hypothesized that the ratio of ITS1 divergence to mitochondrial control region divergence would be approximately 0.5 for each species analyzed (Presa et al. 2002).

The third specific objective of this study was to determine the relationship of the Bermudian endemic anchovy, *Anchoa choerostoma*, to two other species with disparate ranges found in the western central Atlantic, *A. mitchilli* and *A. hepsetus*. It was

hypothesized that *A. choerostoma* would be more closely related to *A. mitchilli* (restricted in range to the U.S. southeast Atlantic coast and the Gulf of Mexico) than to *A. hepsetus* (whose range extends from Canada to the northern coast of South America, as well as throughout the Caribbean). This would lend support to the hypothesis that Bermudian colonists did not originate in the Caribbean, but instead may have originated in populations along the southeastern Atlantic coast of the United States.

The fourth objective of my thesis is to examine the amount of genetic differentiation that separates endemic populations (conspecifics), subspecies, and species in island populations from source populations.

Materials and Method

Species Collection

The species used in this study are tropical and subtropical shorefishes common to both Bermuda and the southeastern United States: *Haemulon aurolineatum*, *H. flavolineatum*, *H. sciurus*, *Holocentrus adscensionis*, *H. rufus*, *Lagodon rhomboides*, and *Lutjanus griseus* (Table 1). Samples were collected from both Bermudian and U.S. populations by seining, spearfishing, rotenone sampling, chevron traps, and hook-and-line. In addition, samples of *Anchoa hepsetus* and *A. mitchilli* were collected in the United States, *A. choerostoma* in Bermuda, and a single putative *A. hepsetus* in Jamaica. *Anchoa choerostoma* is an endemic species of Bermuda, *A. hepsetus* is found throughout the western Atlantic from Canada to South America, and *A. mitchilli* is found in the Gulf of Mexico and along the U.S. Atlantic coastline (Figure 7).

All of the species collected in this study release pelagic eggs and have pelagic larval phases ranging from 13-25 days in the haemulids (McFarland et al. 1985; Lindeman 1986; Lindeman et al. 2001) to 48 days in *Holocentrus adscensionis* (Tyler et al. 1993) (Table 1). These larval phases are of sufficient duration to allow for recruitment to Bermuda under ideal circumstances.

Whole fish were frozen or preserved in 96% ethanol. Voucher specimens were stored in the fish collection at the Virginia Institute of Marine Science under the catalogue numbers 10930 through 10938. For each fish, a small plug of muscle tissue or gill tissue was removed and preserved in DMSO-EDTA storage buffer (0.25M EDTA (pH 8.0), 20% DMSO, saturated with NaCl) for analysis.

Table 1. Pelagic larval durations and number of individuals collected for seven species of shorefishes common to Bermuda and the United States. [1](McFarland et al. 1985; Lindeman 1986) [2](McFarland et al. 1985; Lindeman 1986; Shulman and Bermingham 1995) [3](Tyler et al. 1993) [4](Muncy 1984) [5](Lindeman et al. 2001)

Species	Pelagic larval duration (PLD) in days	Number collected in Bermuda	Number collected in the U.S.	Number collected in Jamaica
Haemulon aurolineatum	13 – 25[1]	20	19 (FL)	-
Haemulon flavolineatum	13-25[2]	18	18 (FL)	
Haemulon sciurus	13-25[1]	20	10 (FL)	
Holocentrus rufus	44[3]	21	9 (FL)	
Holocentrus adscensionis	48[3]	21	9 (FL)	
Lagodon rhomboides	~ 24[4]	21	15 (VA)	
Lutjanus griseus	31-42[5]	24	11 (FL)	
Anchoa choerostoma		13		
Anchoa hepsetus			16	1
Anchoa mitchilli		-	12	-

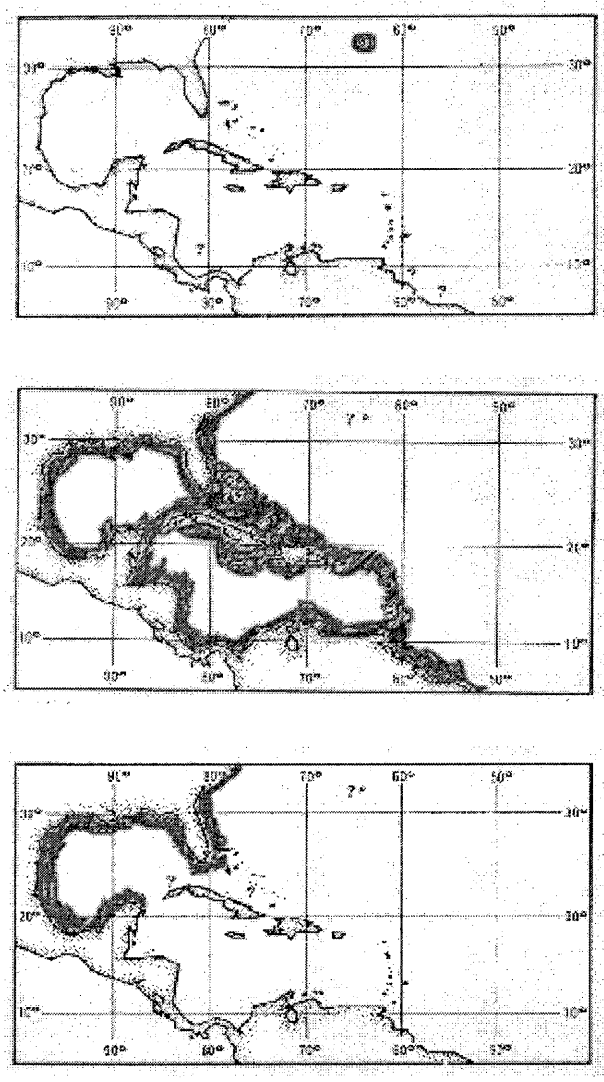


Figure 7. Ranges of three species of western central Atlantic anchovies in the genus *Anchoa*. A: *A. choerostoma*; B: *A. hepsetus*; C: *A. mitchilli*.

DNA Isolation

Genomic DNA was isolated using a standard phenol-chloroform procedure (Sambrook and Russell 2001) with modifications. 0.05 to 0.1 g of tissue was incubated overnight at 37°C in a solution containing 500 µl isolation buffer (50 mM EDTA, 50 mM Tris, 150 mM NaCl, pH 8.0), 60 µl 10% SDS, 10 µl RNAse (10 mg/ml), and 10 µl proteinase K (25 mg/ml).

Following incubation, an equal volume of phenol was added to each tube (approximately 580 µl). The tubes were gently inverted until thoroughly mixed and spun at 14,000g in an Eppendorf 5415 Centrifuge for four minutes at room temperature. The aqueous layer was transferred to a clean microcentrifuge tube using a cut P1000 tip to prevent shearing of the DNA. The extraction was repeated first using an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and again using an equal volume of chloroform:isoamyl alcohol (24:1). DNA was precipitated using 0.04 x volume 5X NaCl and an equal volume of isopropanol, and collected by centrifugation at 14,000g for 18 minutes at 4°C. The isopropanol was decanted and 200 µl of 70% ethanol was added to each tube. This was followed by centrifugation at 14,000g for 5 minutes at 4°C. The ethanol was removed using a pipet and the residual was evaporated using a vacuum concentration system (Savant Speed-Vac 200). Pellets were resuspended in 50 µl of 0.1X filtered TE (pH 8.0) and stored at -20°C. 1.5 µl of each sample was electrophoresed on a 0.8% agarose gel and visualized using ethidium bromide (EtBr) staining under ultraviolet (UV) illumination to ascertain DNA yields.

Mitochondrial DNA

Two sets of primers were used to amplify the mitochondrial control region. For *Haemulon aurolineatum*, *H. flavolineatum*, *H. sciurus*, *Lagodon rhomboides*, *Lutjanus griseus*, *Anchoa choerostoma*, *A. hepsetus*, and *A. mitchilli*, the D-loop primers used for amplification were designed by Cronin (1993): TTGGGTTTCTCGTATGACCG and AGAGCGTCGGTCTTGTAAAAC. For *Holocentrus adscensionis* and *H. rufus*, primers used for amplification were A: TTCCACCTCTAACTCCCAAAGCTAG and G: CGTCGGATCCCATCTTCAGTGTTATGCTT (Lee et al. 1995). PCR reactions were performed using two different systems. Most samples were amplified using puRe *Taq* Ready-To-Go PCR Beads (Amersham Biosciences). Each bead contains stabilizers, BSA, dNTPs, and 2.5 units of puRe *Taq* DNA polymerase. Final reagent concentrations in each reaction were 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, and 200 µM dNTPs. Two 25 µl reactions were performed for each sample. Some samples were amplified using the *Taq* PCR Core Kit (Qiagen). Two 50 µl reactions were performed for each sample, each reaction containing a 1X concentration of 10X Qiagen PCR Buffer, 1.5 mM $MgCl_2$, 200 µl of each dNTP, 0.5 µM of each primer, 2.5 units *Taq* DNA polymerase, and 0.5 µl of DNA template. For samples that proved difficult to amplify, 10 µl of bovine serum albumin (BSA) (1 mg/ml) was added to each 50 µl reaction. The amplification program for both systems contained an initial denaturation at 95°C for 5 minutes, followed by 36 cycles of 94°C for 1 minute, 48-50°C for 1 minute, 72°C for 3 minutes, and a final step of 72°C for 7 minutes. Three µl of each sample were electrophoresed on a 1% agarose gel to ascertain presence of PCR product after EtBr staining.

PCR products were purified using the Wizard PCR Preps DNA Purification System (Promega). The resulting DNA concentration was determined using a DyNA Quant 200 fluorometer (Hoefer). Direct sequencing of the purified PCR products proceeded using a modification of the protocol for simultaneous bi-directional sequencing developed by Li-Cor. The following were combined in a microcentrifuge tube for each sample: 1.5 µl of one infrared (IR)-labeled primer (1CD-loop (IR800) or 2CD-loop(IR700)), double-stranded DNA template (volume determined by the formula: X µl template = 0.05/(0.85 x concentration of DNA in µg/ml)), and PCR H_20 to a final volume of 17 µl. Four µl of the template/primer mix was added to each of four 0.2 ml tubes corresponding to that sample. One µl of ThermoSequenase Primer Cycle Sequencing Kit (Amersham Biosciences) reagent mix containing the appropriate dideoxynucleotide terminator was added to each tube - A, G, C, or T. Approximately 10 µl of silicon oil was added to each tube to prevent evaporation of the reaction. The same procedure was repeated for each sample using the second IR-labeled primer. The reaction program contained an initial denaturation at 95°C for 5 minutes, followed by 30 cycles of 92°C for 30 seconds, 58°C for 30 seconds, 70°C for 30 seconds. After cycling, 3 µl of stop solution containing formamide was added under the silicon oil. Each set of reactions (one set for each primer) was denatured for 3 minutes at 92°C prior to being loaded on a 66 cm, 0.2 mm KB^+ polyacrylamide gel in a Li-Cor Global Edition IR^2 DNA Sequencer. The gels were made by adding 267 µl of 10% ammonium persulfate and 24 µl of Temed to 40 ml of 3.7% KB^+ polyacrylamide gel matrix (Li-Cor) which allowed for polymerization. The two sets of reactions were loaded 5-10 minutes apart. Electrophoresis conditions followed Li-Cor recommendations.

For four individuals of *Holocentrus adscensionis*, four *H. rufus*, and one *Lagodon rhomboides*, direct sequencing proved to be very difficult. The control region from these individuals was cloned and bidirectionally sequenced using protocols described in the following section.

Nuclear DNA

The first nuclear internal transcribed spacer region (ITS1) of the ribosomal RNA gene complex was originally amplified using the primers ITS3: TATGCTTAAATTCAG CGGGT and ITS5: CGTAGGTGAACCTGCGGAAGG (Goggin 1994). However, alignments revealed that the resulting ITS sequences were almost identical across all species examined. The sequences were compared with those deposited in GenBank through NCBI BLAST search, and it was discovered that ITS had been amplified from protists within the fish tissue, not from the fish themselves.

I subsequently developed eight sets of primers using 18S and 5.8S sequences from fishes and the amphibian genus *Xenopus* in GenBank and the computer program MacVector (Kaufman et al. 1994). A series of test amplifications were completed to determine which primer sets were most successful for each species examined. For the purposes of sequencing, samples of *Haemulon aurolineatum*, *H. flavolineatum*, *H. sciurus*, *Lagodon rhomboides*, and *Lutjanus griseus* were amplified using the primers 18SF2: TAGTGAGGTCCTCGGAT and 5.8SR2: GTGCGTTCGAAR KGTCGATGATCAAT. Samples of *Holocentrus adscensionis*, *H. rufus*, *Anchoa choerostoma*, *A. hepsetus*, and *A. mitchilli* were amplified using the primers X18SF: CTTGACTATCTAGAGGAAGT and X28SR: ATATGCTTA AATTCAGCGGG.

Reactions were performed using the *Taq* PCR Core Kit (Qiagen). One 10 µl reaction was performed for each sample and each reaction contained a 1X concentration of 10X Qiagen PCR Buffer, 1.5 mM $MgCl_2$, 200 µl of each dNTP, 0.5 µM of each primer, 2.5 units *Taq* DNA polymerase, and 0.1 µl of DNA template. For samples that proved difficult to amplify, either 2 µl of BSA (1 mg/ml) or 2 µl of 5X Q-Solution (Qiagen) were added to each 10 µl reaction. The amplification program contained an initial denaturation at 94°C for 5 minutes, followed by 35 cycles of 94°C for 1 minute, 45-49°C for 1 minute, 72°C for 3 minutes, followed by a final step of 72°C for 7 minutes. Three µl of each sample were run on a 1% agarose gel to ascertain presence of PCR product. Bands were visualized using EtBr staining.

PCR products were ligated into the pCR 4-TOPO vector using the TOPO TA Cloning Kit for Sequencing and the protocols listed in the TOPO TA Cloning Kit for Sequencing Manual Version J (Invitrogen Life Technologies). Transformation of the ligation reactions into Top 10 Chemically Competent *E. coli* cells was performed following the TOPO TA Cloning Kit Manual Version J. Transformed bacterial colonies were restreaked into a grid on a new LB-amp plate, incubated overnight at 37°C, and subsequently "quickscreened" for the presence of inserts. A sterile toothpick was used to resuspend a colony patch in 30 µl of STE (100 mM NaCl, 20 mM Tris-HCl (pH 7.5), 10 mM EDTA (pH8.0)). Thirty µl of phenol:chloroform:isoamyl alcohol (25:24:1) was added and the tubes were vortexed and centrifuged at room temperature for 5 minutes at 14,000g. One µl of RNAse and 2 µl of loading dye were added to 15 µl of sample supernatant. Samples were electrophoresed on a 1% agarose gel and visualized using EtBr staining.

Colonies found to contain inserts were grown overnight (between 12 and 13 hours) in 3 ml of LB broth containing 6 μl ampicillin (50 mg/ml). Isolation of the transformed plasmid from the bacterial genomic DNA was performed using the QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's protocols. A DyNA Quant 200 fluorometer was used to determine the resulting concentration of DNA. Sequencing proceeded following the simultaneous bi-directional sequencing protocol developed by Li-Cor. The volume of template DNA added to each reaction was determined using the following formula: X μl template = 0.25/(0.31 x concentration of DNA in μg/ml). The reaction program contained an initial denaturation at 95°C for 5 minutes, followed by 30 cycles of 92°C for 30 seconds, 54°C for 30 seconds, 70°C for 30 seconds.

After cycling, 3 μl of stop solution containing formamide were added to each reaction. Reactions were denatured for 3 minutes at 92°C prior to being loaded on a 66 cm, 0.2 mm KB^+ polyacrylamide gel in a Li-Cor Global Edition IR^2 DNA Sequencer. Polymerization of the gel and electrophoresis conditions were the same as those previously described in the section on mitochondrial DNA.

Between one and three clones were sequenced for each individual (Table 2). As ITS is a multi-copy nuclear marker, several different allelic forms could be isolated from a single individual.

Data Analysis

For both mitochondrial and nuclear sequences, v. 2 standard chromatogram file (SCF) curves were generated using the e-Seq DNA Sequencing and Analysis Software v. 2.0 (Li-Cor) and imported into the computer program Sequencher 4.1 (GeneCodes

Corporation) to be edited and assembled. Separate alignments of the sequences for each species were performed using the ClustalW Multiple Sequence Alignment Program version 1.7 (Thompson 1994) in the computer program MacVector 7.0 (Kaufman et al. 1994). Different gap penalties were used for different species to optimize each alignment independently. For the control region alignments, in both the pairwise and multiple alignment phases, the open gap penalties and extend gap penalties used were respectively 10 and 5 in *Lagodon rhomboides*, 8 and 3 in *Lutjanus griseus*, *Holocentrus rufus*, *Haemulon aurolineatum*, *H. flavolineatum*, *H. sciurus*, and the *Anchoa* spp., and 7 and 1 in *Holocentrus adscensionis*. Delay divergence values were 40%, and transitions were unweighted. For the ITS1 alignments, in both the pairwise and multiple alignment phases, the open gap penalties and extend gap penalties used were respectively 10 and 5 in *Haemulon aurolineatum* and *Lagodon rhomboides*, 8 and 3 in *Haemulon flavolineatum*, *H. sciurus*, *Holocentrus adscensionis*, and *Lutjanus griseus*, and 7 and 1 in *Holocentrus rufus* and the *Anchoa* spp.

Sequence variation was analyzed in up to 16 individuals (eight each from Bermuda and the United States) of each species. The computer program DNASequencePolymorphism 3.51 (DNAsp) (Rozas and Rozas 1999) was used to calculate haplotype (gene) diversity, the probability of sampling different haplotypes when two individuals are drawn from the same population, and the number of haplotypes (alleles) within each sample. The computer program MEGA 2.1 (Kumar et al. 2001) was used to calculate the number of conserved sites, the number of variable sites, the mean distance within groups or mean nucleotide sequence diversity (π) (the average number of differences per site between sequences within the same population), mean distance

between groups or mean nucleotide sequence divergence (the average number of differences per site between sequences in two separate populations), net distance between groups or net mean nucleotide sequence divergence (δ) (corrected for the amount of variation within each population), and distance based neighbor-joining trees. All values were calculated using the Tamura-Nei model of evolution.

In addition, MEGA was used to determine the number of transitions, transversions, and the transition/transversion ratio for both the entire mitochondrial control region and an approximately 300 basepair subsection of the control region at the 5' end near the tRNA-proline gene. A recent correlation has been noted between these two transition/transversion ratios and the rate of sequence evolution of the mitochondrial control region exhibited by a given species (McMillan and Palumbi 1997).

The computer program Arlequin (Schneider et al. 2000) was used to calculate Fst between Bermudian and U.S. populations (a measure of the reduction in heterozygosity due to genetic drift, a fixation index widely used to indicate population structure), Analysis of Molecular Variance, and to calculate branch lengths for minimum spanning networks.

To determine if the species examined all colonized Bermuda simultaneously approximately 18,000 years ago, mtDNA control region rates of sequence evolution of 3.6 ± 0.46% per million years (Donaldson and Wilson 1999) and 11.5% per million years (Vigilant et al. 1991) were used to calculate the amount of time since divergence between the Bermudian and U.S. populations. These rates were chosen based upon the transition/transversion ratios exhibited by each species. The rate of sequence evolution of ITS1 is approximately one half that of the control region (Presa et al. 2001). In order

to calculate time since divergence based on ITS1 data, the rates of sequence evolution of the control region were divided by 2.

One-way ANOVAs were performed in the computer program MINITAB to determine if haplotype diversities, gene diversities, and mean nucleotide sequence diversities for both the mitochondrial control region and the nuclear ITS1 region were significantly different between Bermuda and the U.S. In addition, one-way ANOVAs were performed to test the significance of net mean nucleotide sequence divergence values for the control region and ITS1 versus zero, and to compare the estimated times of divergence between the control region and ITS1.

Results

Sixteen individuals of *Haemulon aurolineatum*, *H. flavolineatum*, *H. sciurus*, *Holocentrus adscensionis*, *H. rufus*, and *Lagodon rhomboides* were sequenced for analysis, eight from Bermuda and eight from the United States. For *Lutjanus griseus*, eight individuals were analyzed from Bermuda and seven from the United States. For the *Anchoa* species, eight *A. choerostoma* were sequenced from Bermuda, eight each of *A. hepsetus* and *A. mitchilli* were sequenced from the U.S., and one *A. hepsetus* was sequenced from Jamaica.

Mitochondrial DNA

The control region ranged in length from 860 basepairs (bp) in *Holocentrus rufus* to 1,355 bp in *Holocentrus adscensionis* (Table 2). The number of variable positions within the control region ranged from 39 in *Haemulon aurolineatum* to 231 among the *Anchoa* species. Seven of the species demonstrated no intraspecific length variation, while an 83 bp repeat motif was observed at the 3' end of the control region in *Holocentrus adscensionis*. Among the eight U.S. samples, the number of repeats ranged from zero to six, while the number of repeats found in the eight Bermudian samples ranged from two to five (Table 3).

The transition/transversion ratios for the entire control region ranged from 7:2 in *Haemulon aurolineatum* to 35:3 in *Lagodon rhomboides* (Table 4). When the 300 bp subsection of the control region on the 5' end adjacent to the tRNA proline gene was considered separately, values ranged from 5:1 in *Haemulon aurolineatum* to 19:2 in *Holocentrus rufus* (Table 4). The three haemulid species had similar ratios of

Table 2. Length of the control region sequence and number of variable sites for each species. N=16 for all species except *Lutjanus griseus* (N=15) and *Anchoa* spp. (N=25).

Species	**Length of control region (bp)**	**Number of variable sites**
Haemulon aurolineatum	947	39
Haemulon flavolineatum	965	113
Haemulon sciurus	957	125
Holocentrus adscensionis	863-1355	97
Holocentrus rufus	860	148
Lagodon rhomboides	966	123
Lutjanus griseus	869	43
Anchoa species	1038	231

Table 3. The number of 3' end repeats in the control region per sample from each location for the species *Holocentrus adscensionis*.

Location	Sample Number	Number of repeats
Bermuda	1	2
Bermuda	2	3
Bermuda	3	5
Bermuda	4	4
Bermuda	6	2
Bermuda	9	5
Bermuda	10	2
Bermuda	11	2
U.S.	1	0
U.S.	2	3
U.S.	3_1 (cloned)	5
U.S.	4_1 (cloned)	3
U.S.	5	2
U.S.	6_4 (cloned)	5
U.S.	7_6 (cloned)	6
U.S.	9	3

Table 4. Transition/tranversion rations for the entire control region and for an approximately 300 bp section on the 5' tRNA proline end of the control region for each species.

Species	**Entire control region Si/Sv**	**5' tRNA Proline end Si/Sv**
Haemulon aurolineatum	7:2 (3.5)	5:1 (5.0)
Haemulon flavolineatum	12:5 (2.4)	7:3 (2.3)
Haemulon sciurus	24:8 (3.0)	13:5 (2.6)
Holocentrus adscensionis	17:1 (17.0)	9:0 (~9.0)
Holocentrus rufus	34:3 (11.3)	19:2 (9.5)
Lagodon rhomboides	35:3 (11.7)	17:2 (8.5)
Lutjanus griseus	12:2 (6.0)	9:1 (9.0)

approximately 3:1 for both the entire control region and for the 300 bp subsection while the transition/transversion ratios for the remaining species were similar with a value of approximately 10:1.

Haplotype diversity (*h*) ranged from 0.464 in *Anchoa choerostoma* to 1.00 in several different collections, both in the U.S. and Bermuda (Table 5). Haplotypes were shared between Bermuda and U.S. collections of only one species, *Haemulon flavolineatum*. Specifically, three haplotypes were common to both populations, representing a total of seven individuals. The haplotype diversity and number of haplotypes observed in Bermuda collections was less than or equal to that found in the U.S. collections for all comparisons, although this trend was not significant ($F=2.40$, $p=0.142$). It should be noted that for both Bermudian and U.S. collections of *Lutjanus griseus*, haplotype diversity was 1.00 with eight individuals sequenced from Bermuda while only seven were sequenced from the U.S. The eighth sample from the U.S. population was revealed genetically to be *Lutjanus campechanus* (red snapper) after sequencing and comparison with sequences in GenBank through NCBI BLAST search. There is no voucher specimen of this individual, but the sequence was removed from subsequent analyses due to likely misidentification of the sample.

For the *Anchoa* species, the haplotype diversity and number of haplotypes was smallest in the endemic Bermudian population, *A. choerostoma*, with values of 0.464 and 3, respectively. Both haplotype diversity and numbers of haplotypes increased with increasing population range (Figure 7) in the remaining two species, *A. mitchilli* and *A. hepsetus*. *A. mitchilli* had a haplotype diversity of 0.964 with seven haplotypes while *A. hepsetus* had a haplotype diversity of 1.00 with eight haplotypes.

Table 5. Number of control region haplotypes, haplotype diversity (h), and mean nucleotide sequence diversity (π) for each species by location. Numbers in parentheses indicate the total number of individuals sequenced at that location.

Species	Location	Number of haplotypes	Haplotype diversity (h)	Mean nucleotide sequence diversity (π)
Haemulon aurolineatum	Bermuda	5 (8)	0.857 ± 0.108	0.0015 ± 0.0007
	U.S.	8 (8)	1.00 ± 0.063	0.0136 ± 0.0023
Haemulon flavolineatum	Bermuda	7 (8)	0.964 ± 0.077	0.0055 ± 0.0015
	U.S.	8 (8)	1.00 ± 0.063	0.0332 ± 0.0034
Haemulon sciurus	Bermuda	7 (8)	0.964 ± 0.077	0.0066 ± 0.0019
	U.S.	8 (8)	1.00 ± 0.063	0.0578 ± 0.0055
Holocentrus adscensionis	Bermuda	8 (8)	1.00 ± 0.052	0.01909 ± 0.0026
	U.S.	8 (8)	1.00 ± 0.063	0.01785 ± 0.0026
Holocentrus rufus	Bermuda	8 (8)	1.00 ± 0.063	0.0401 ± 0.0039
	U.S.	8 (8)	1.00 ± 0.063	0.0445 ± 0.0043
Lagodon rhomboides	Bermuda	7 (8)	0.964 ± 0.077	0.0391 ± 0.0042
	U.S.	8 (8)	1.00 ± 0.063	0.0447 ± 0.0045
Lutjanus griseus	Bermuda	8 (8)	1.00 ± 0.063	0.0146 ± 0.0028
	U.S.	7 (7)	1.00 ± 0.076	0.0171 ± 0.0028
Anchoa species	*A. choerostoma*	3 (8)	0.464 ± 0.200	0.0016 ± 0.0007
	A. hepsetus	8 (8)	1.00 ± 0.004	0.0088 ± 0.0018
	A. mitchilli	7 (8)	0.964 ± 0.077	0.0039 ± 0.0013

Mean nucleotide sequence diversity (π) of the Bermudian collections ranged from 0.0015 in *Haemulon aurolineatum* to 0.0401 in *Holocentrus rufus* (Table 5). The U.S. collections had values ranging from 0.0039 in *Anchoa mitchilli* to 0.0578 in *Haemulon sciurus* (Table 5). The diversity of the three haemulid species was an order of magnitude smaller in Bermudian collections (0.0015 – 0.0066) than in U.S collections (0.0136 – 0.0578). While not an order of magnitude in difference, the same pattern of smaller diversity in Bermudian collections holds true for all other species examined except *Holocentrus adscensionis*. This species showed a slightly larger diversity in the Bermuda collection, 0.01909, than in the U.S. collection, 0.01785. The overall trend of lower mean nucleotide sequence diversity in Bermuda compared to the United States is not significant, however, as revealed by one-way ANOVA ($F=1.62$, $p=0.222$).

For the *Anchoa* species, diversity was smallest in the endemic Bermudian anchovy, *A. choerostoma*, with a value of 0.0016. This species of *Anchoa* also had the smallest haplotype diversity. Mean nucleotide sequence diversity increased with increasing population range (and haplotype diversity) for the remaining two species, *A. mitchilli* with a value of 0.0039 and *A. hepsetus* with a value of 0.088.

Analyses within each collection revealed an overall trend of reduced genetic diversity in Bermuda (although not significant), while analyses between each collection revealed divergence estimates varying by an order of magnitude. The mean nucleotide sequence divergence between Bermudian and U.S. collections ranged from 0.0114 in *Haemulon aurolineatum* to 0.0488 in *Holocentrus rufus* (Table 6). The net mean nucleotide sequence divergence between Bermudian and U.S. collections ranged from 0.000236 in *Lutjanus griseus* to 0.007438 in *Haemulon sciurus* (Table 6).

Table 6. Mean nucleotide sequence divergence and net mean nucleotide sequence divergence between Bermudian and U.S. populations of each species for the control region.

Species	Mean nucleotide sequence divergence	Net mean nucleotide sequence divergence
Haemulon aurolineatum	0.0114 ± 0.0022	0.003961 ± 0.001316
Haemulon flavolineatum	0.0190 ± 0.0020	0.000348 ± 0.000163
Haemulon sciurus	0.0400 ± 0.0037	0.007438 ± 0.001212
Holocentrus adscensionis	0.0177 ± 0.0022	0.000731 ± 0.000315
Holocentrus rufus	0.0488 ± 0.0042	0.006492 ± 0.001590
Lagodon rhomboides	0.0415 ± 0.0042	0.000440 ± 0.000470
Lutjanus griseus	0.0161 ± 0.0027	0.000236 ± 0.000614

Mean nucleotide sequence divergence among the *Anchoa* species ranged from 0.053 between *A. hepsetus* (Jamaica) and *A. mitchilli* to 0.165 between *A. choerostoma* and *A. hepsetus* (U.S.) (Table 7). Net mean nucleotide sequence divergence ranged from 0.051 between *A. hepsetus* (Jamaica) and *A. mitchilli* to 0.160 between *A. choerostoma* and *A. hepsetus* (U.S.). The endemic Bermudian anchovy, *A. choerostoma*, has its lowest mean nucleotide sequence divergence (0.123) and lowest net mean nucleotide sequence divergence (0.120) with *A. mitchilli.*

Fst values were calculated in order to determine if population structure existed; these values ranged from -0.00832 (0) in *Lagodon rhomboides* to 0.96829 among the *Anchoa* species (Table 8). Fst values were significant between Bermudian and U.S. collections of *Haemulon aurolineatum* (0.34756, $p<0.01$), *Haemulon sciurus* (0.19695, $p=0.002$), and *Holocentrus rufus* (0.12565, $p=0.022$). For these three species, AMOVA revealed that 34.76%, 19.69%, and 12.57% of the variance was found among populations, respectively. The value of Fst between the *Anchoa* species was also significant (0.96829, $p<0.01$), with 96.83% of the variance found among the species. Other comparisons were not significant.

Neighbor-joining trees based on mitochondrial control region sequences for each species and minimum spanning networks for species with significant Fst values are shown in Figures 8 – 15. The Bermudian samples of *Haemulon aurolineatum* cluster together in a clade separate from the U.S. samples (Figure 8A). The minimum spanning network for this species reveals two of the Bermudian haplotypes are shared between five individuals and the other three haplotypes are only 1 or 2 basepairs different. This clade of Bermudian samples is connected by a single basepair change to a more structured U.S.

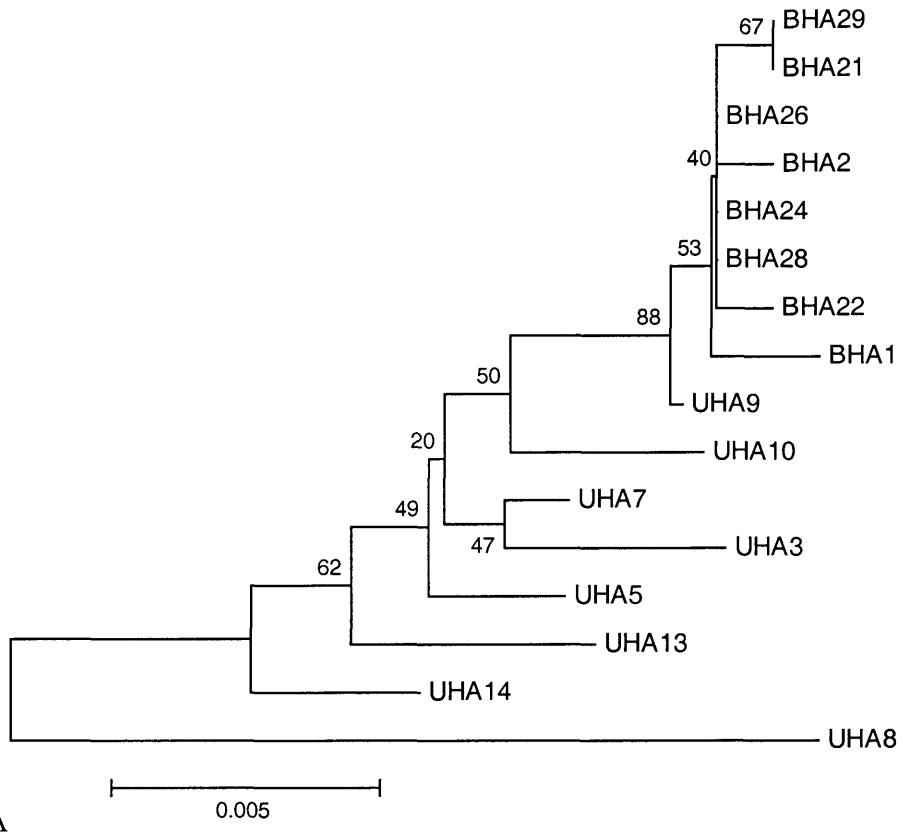
Table 7. Mean nucleotide sequence divergence (below the diagonal) and net mean nucleotide sequence divergence (above the diagonal) between the species of *Anchoa* for the control region.

	A. mitchilli	*A. hepsetus*	*A. hepsetus (Jamaica)*	*A. choerostoma*
A. mitchilli	--	0.147 ± 0.013	0.051 ± 0.007	0.120 ± 0.012
A. hepsetus	0.153 ± 0.013	--	0.142 ± 0.013	0.160 ± 0.013
A. hepsetus (Jamaica)	0.053 ± 0.007	0.146 ± 0.013	--	0.122 ± 0.012
A. choerostoma	0.123 ± 0.012	0.165 ± 0.013	0.123 ± 0.012	--

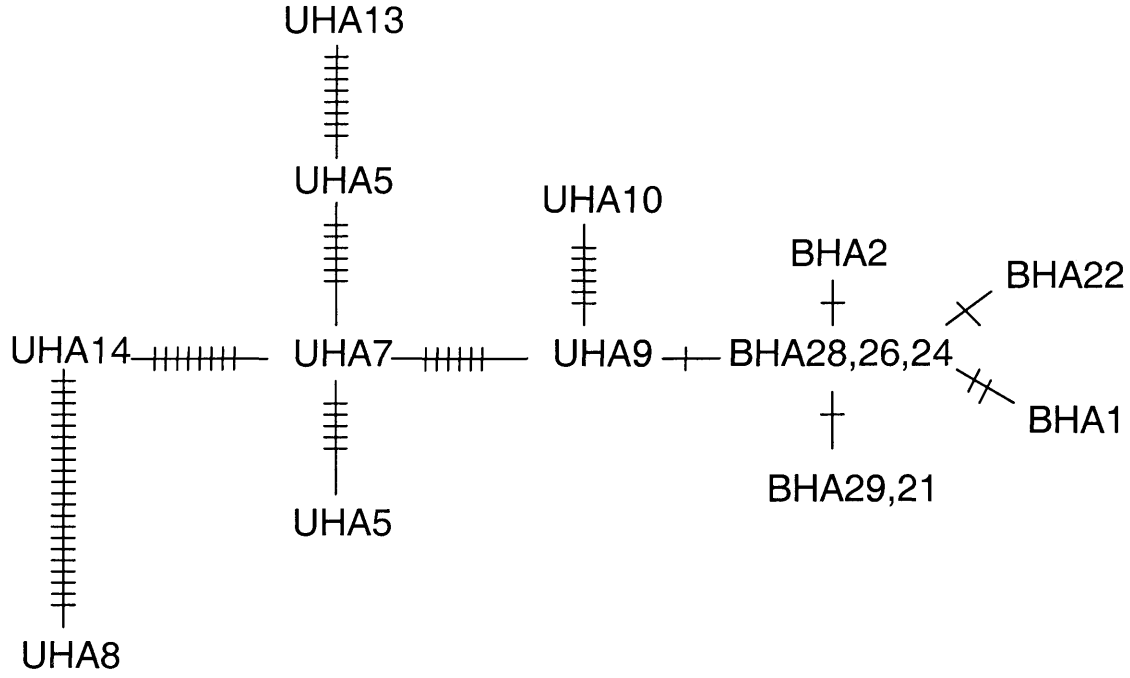
Table 8. Fst and AMOVA results based upon control region sequences for each species.

Species	**Fst**	**AMOVA variance among populations**	**AMOVA variance within populations**
Haemulon aurolineatum	0.34756 (p=0.000)	34.76%	65.24%
Haemulon flavolineatum	-0.01987 (p=0.893)	-1.99%	101.99%
Haemulon sciurus	0.19695 (p=0.002)	19.69%	80.31%
Holocentrus adscensionis	-0.05366 (p=6310	-5.37%	105.37%
Holocentrus rufus	0.12565 (p=0.022)	12.57%	87.43%
Lagodon rhomboides	-0.00832 (p=0.396)	-0.83%	100.83%
Lutjanus griseus	0.01620 (p=0.305)	1.62%	98.38%
Anchoa spp.	0.96829 (p=0.000)	96.83%	3.17%

Figure 8. A) Neighbor-joining tree of control region sequences from *Haemulon aurolineatum* with bootstrap support values. B) Minimum spanning network of Bermudian and U.S. haplotypes. BHA = Bermudian samples; UHA = U.S. samples.


67
BHA29
BHA21
BHA26
40
BHA2
BHA24
53
BHA28
BHA22
88
BHA1
50
UHA9
20
UHA10
UHA7
49
47
UHA3
62
UHA5
UHA13
UHA14
UHA8
0.005

A


UHA13
UHA5
UHA10
BHA2
BHA22
UHA14
UHA7
UHA9
BHA28,26,24
BHA1
BHA29,21
UHA5
UHA8

B

Figure 9. Neighbor-joining tree of control region sequences from *Haemulon flavolineatum* with bootstrap support values. BHF = Bermudian samples; UHF = U.S. samples.

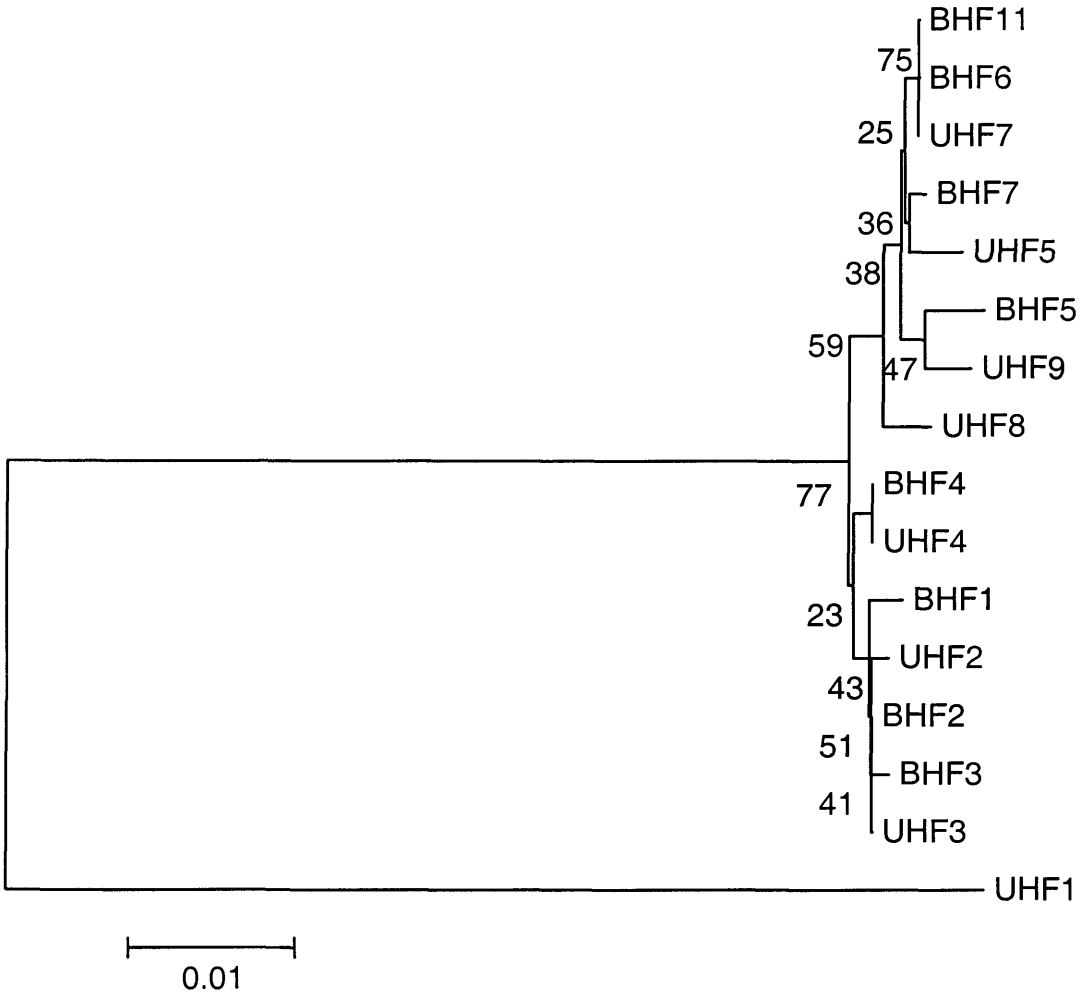
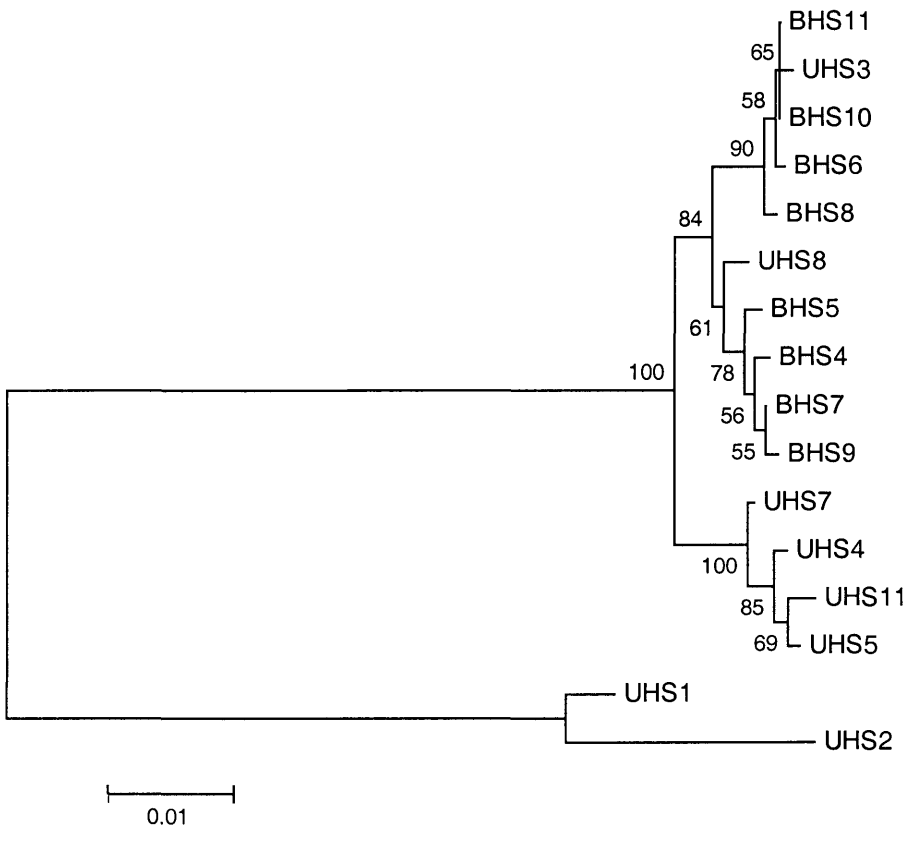
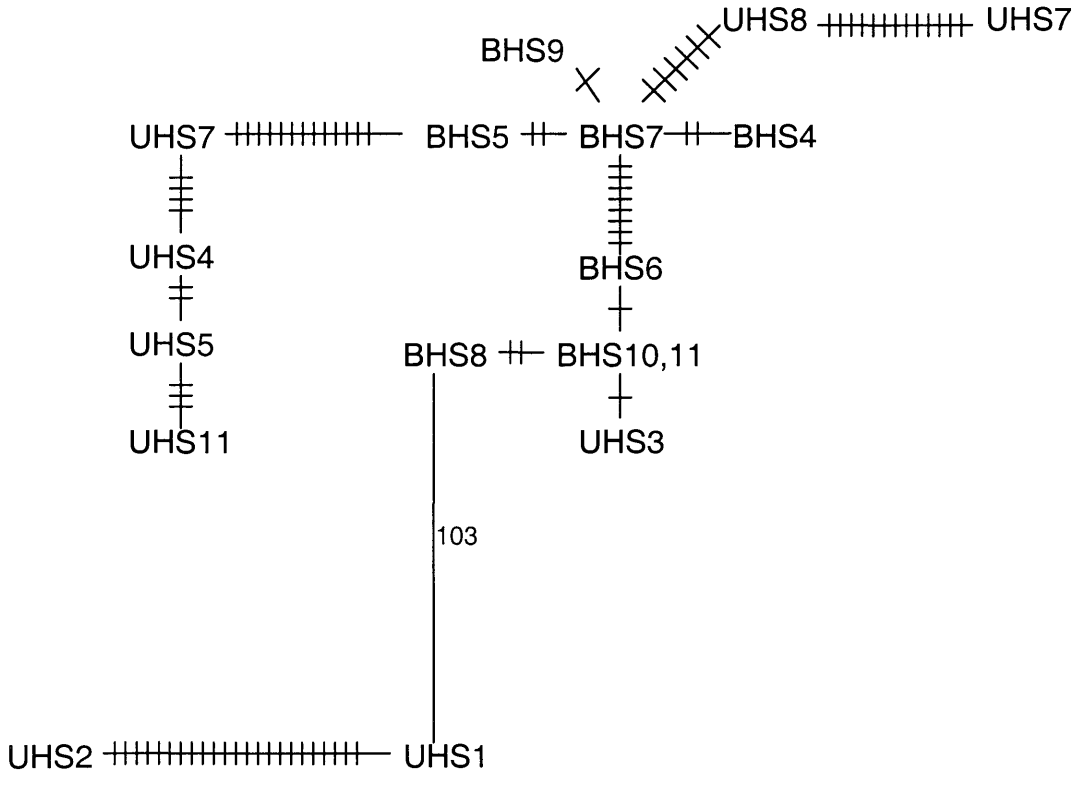

BHF11
75
BHF6
25
UHF7
BHF7
36
UHF5
38
BHF5
59
47
UHF9
UHF8
77
BHF4
UHF4
23
BHF1
UHF2
43
BHF2
51
BHF3
41
UHF3
UHF1
0.01

Figure 10. A) Neighbor-joining tree of control region sequences from *Haemulon sciurus* with bootstrap support values. B) Minimum spanning network of Bermudian and U.S. haplotypes. BHS = Bermudian samples; UHS = U.S. samples.


BHS11
65
UHS3
58
BHS10
90
BHS6
BHS8
84
UHS8
BHS5
61
BHS4
100
78
BHS7
56
55
BHS9
UHS7
UHS4
100
85
UHS11
69
UHS5
UHS1
UHS2
0.01

A


UHS8
UHS7
BHS9
UHS7
BHS5
BHS7
BHS4
UHS4
BHS6
UHS5
BHS8
BHS10,11
UHS11
UHS3
103
UHS2
UHS1

B

Figure 11. Neighbor-joining tree of control region sequences from *Holocentrus adscensionis* with bootstrap support values. BHOA = Bermudian samples; UHOA = U.S. samples.

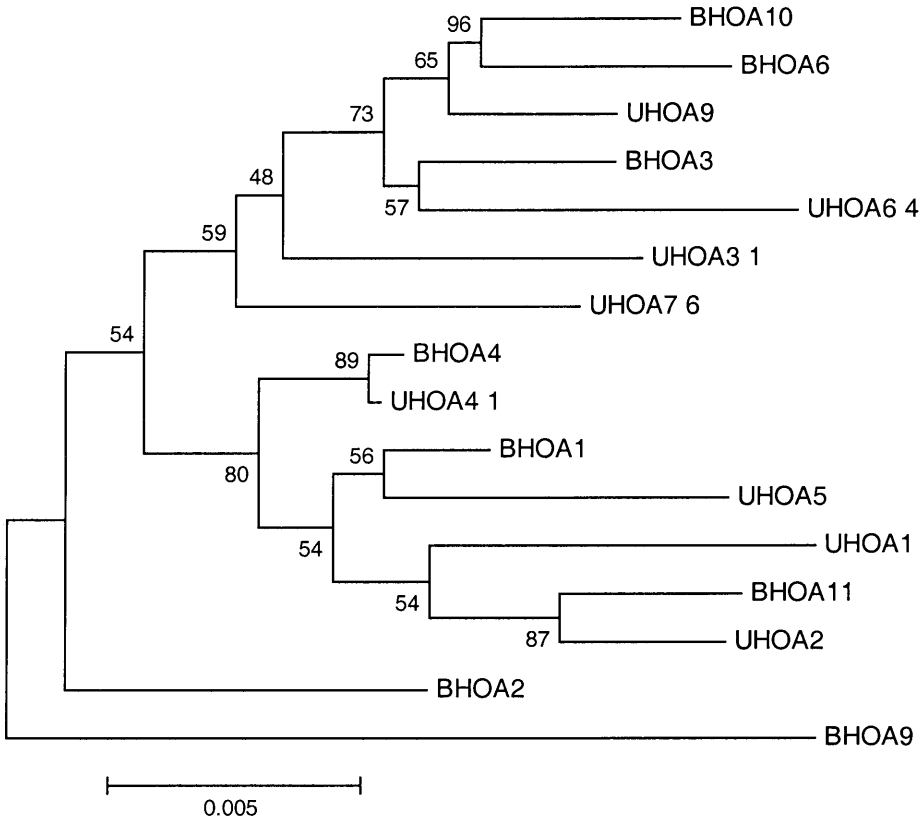
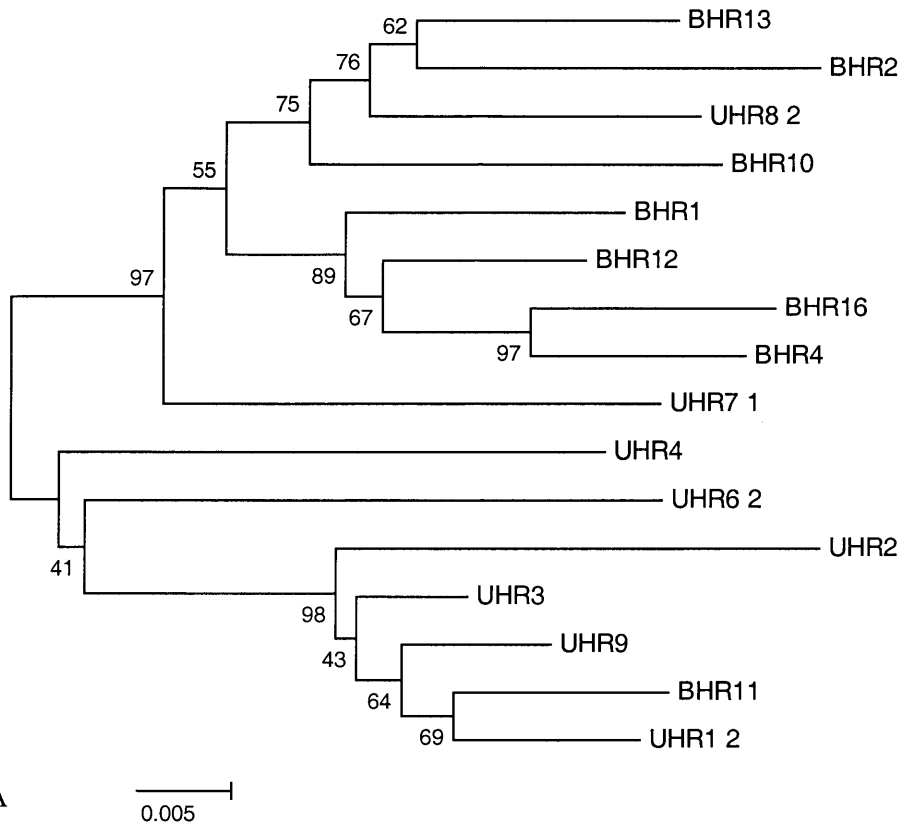
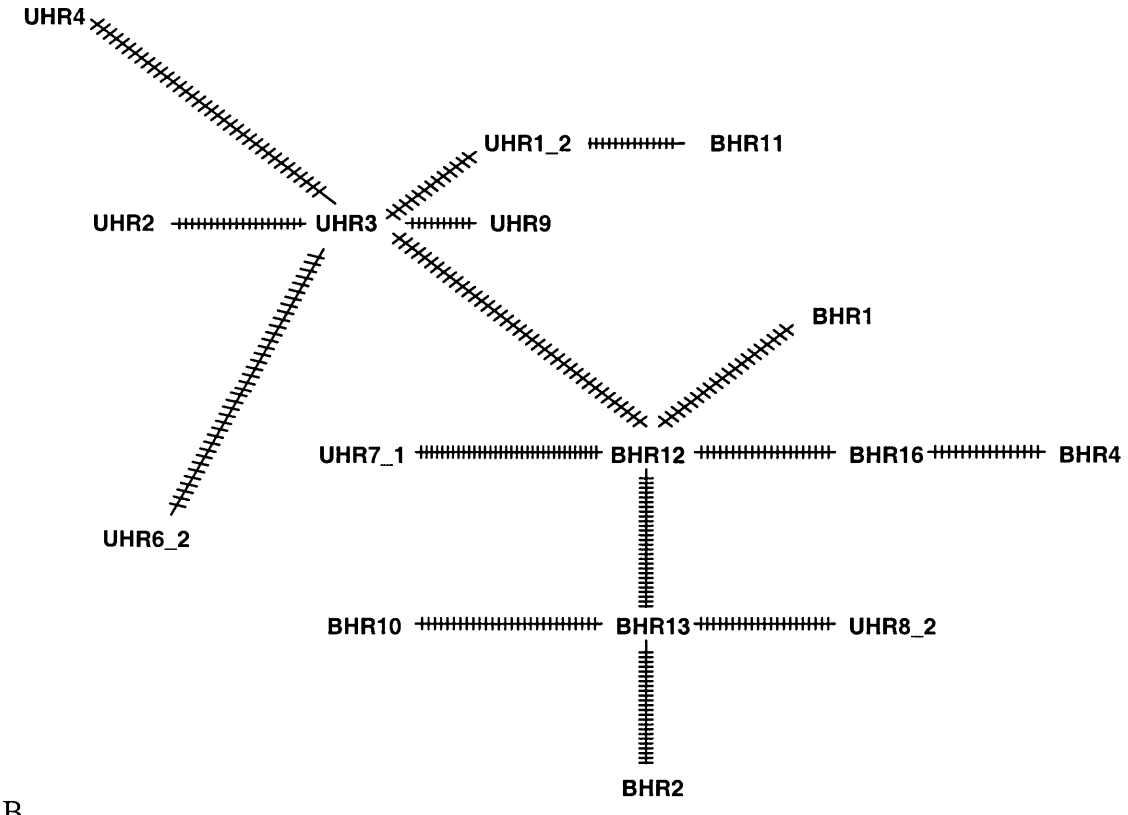

BHOA10
BHOA6
UHOA9
BHOA3
UHOA6 4
UHOA3 1
UHOA7 6
BHOA4
UHOA4 1
BHOA1
UHOA5
UHOA1
BHOA11
UHOA2
BHOA2
BHOA9
96
65
73
48
57
59
54
89
80
56
54
54
87
0.005

Figure 12. A) Neighbor-joining tree of control region sequences from *Holocentrus rufus* with bootstrap support values. B) Minimum spanning network of Bermudian and U.S. haplotypes. BHR = Bermudian samples; UHR = U.S. samples.


BHR13
BHR2
UHR8 2
BHR10
BHR1
BHR12
BHR16
BHR4
UHR7 1
UHR4
UHR6 2
UHR2
UHR3
UHR9
BHR11
UHR1 2
62
76
75
55
97
89
67
97
41
98
43
64
69

A

0.005


UHR4
UHR1_2
BHR11
UHR2
UHR3
UHR9
BHR1
UHR7_1
BHR12
BHR16
BHR4
UHR6_2
BHR10
BHR13
UHR8_2
BHR2

B

Figure 13. Neighbor-joining tree of control region sequences from *Lagodon rhomboides* with bootstrap support values. BLR = Bermudian samples; ULR = U.S. samples.

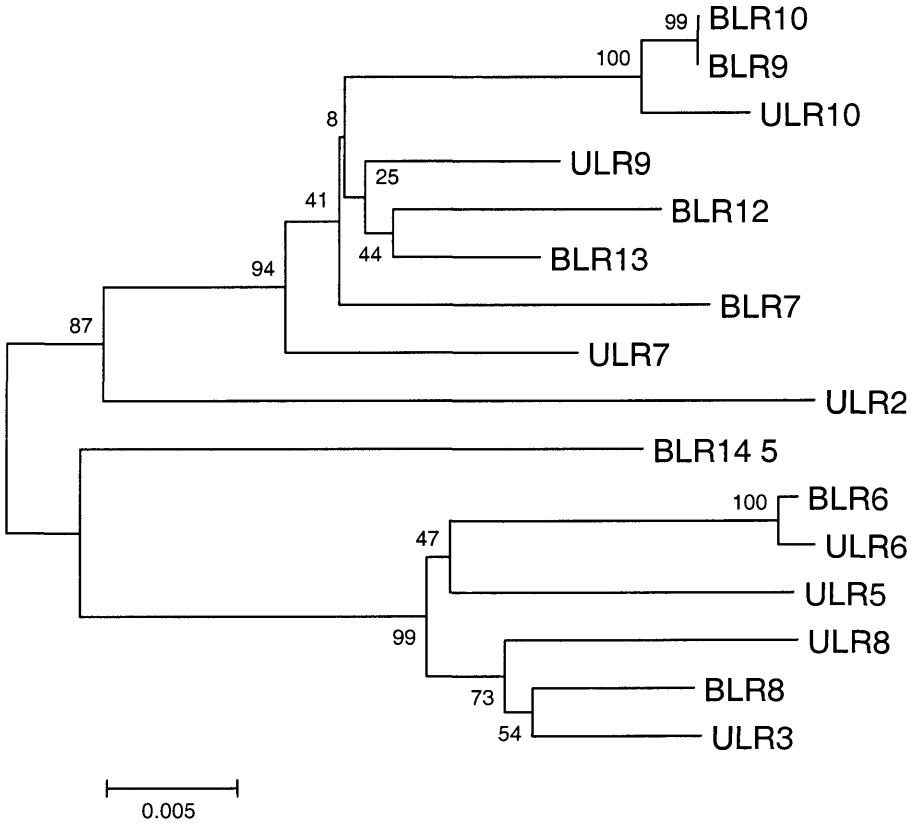

99
BLR10
100
BLR9
ULR10
8
ULR9
25
41
BLR12
44
BLR13
94
BLR7
87
ULR7
ULR2
BLR14 5
100
BLR6
47
ULR6
ULR5
99
ULR8
73
BLR8
54
ULR3
0.005

Figure 14. Neighbor-joining tree of control region sequences from *Lutjanus griseus* with bootstrap support values. BLG = Bermudian samples; ULG = U.S. samples.

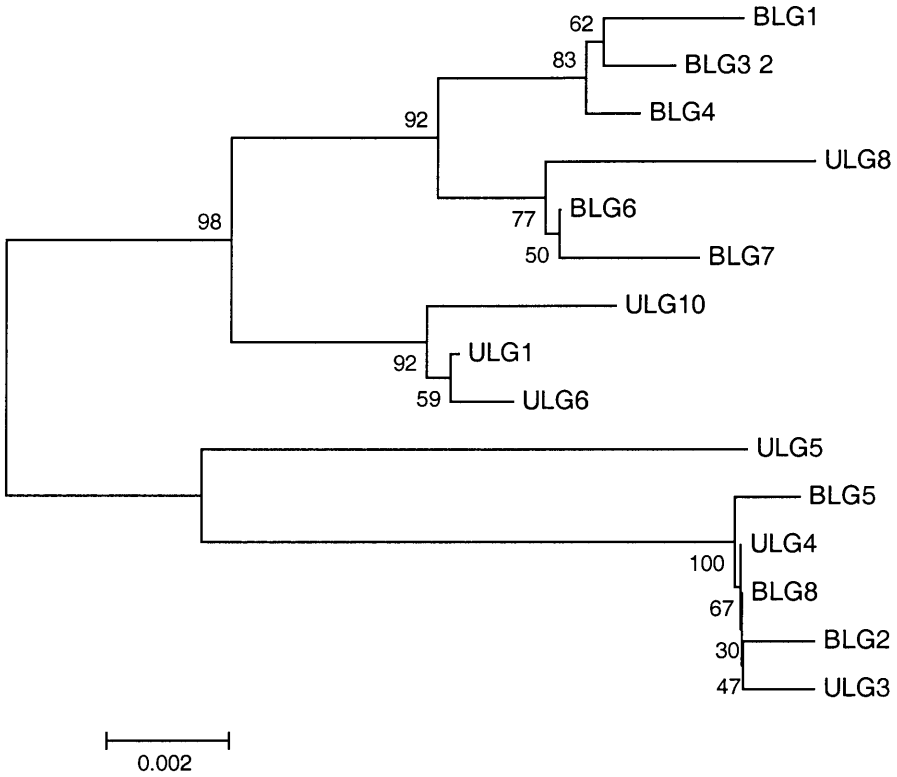
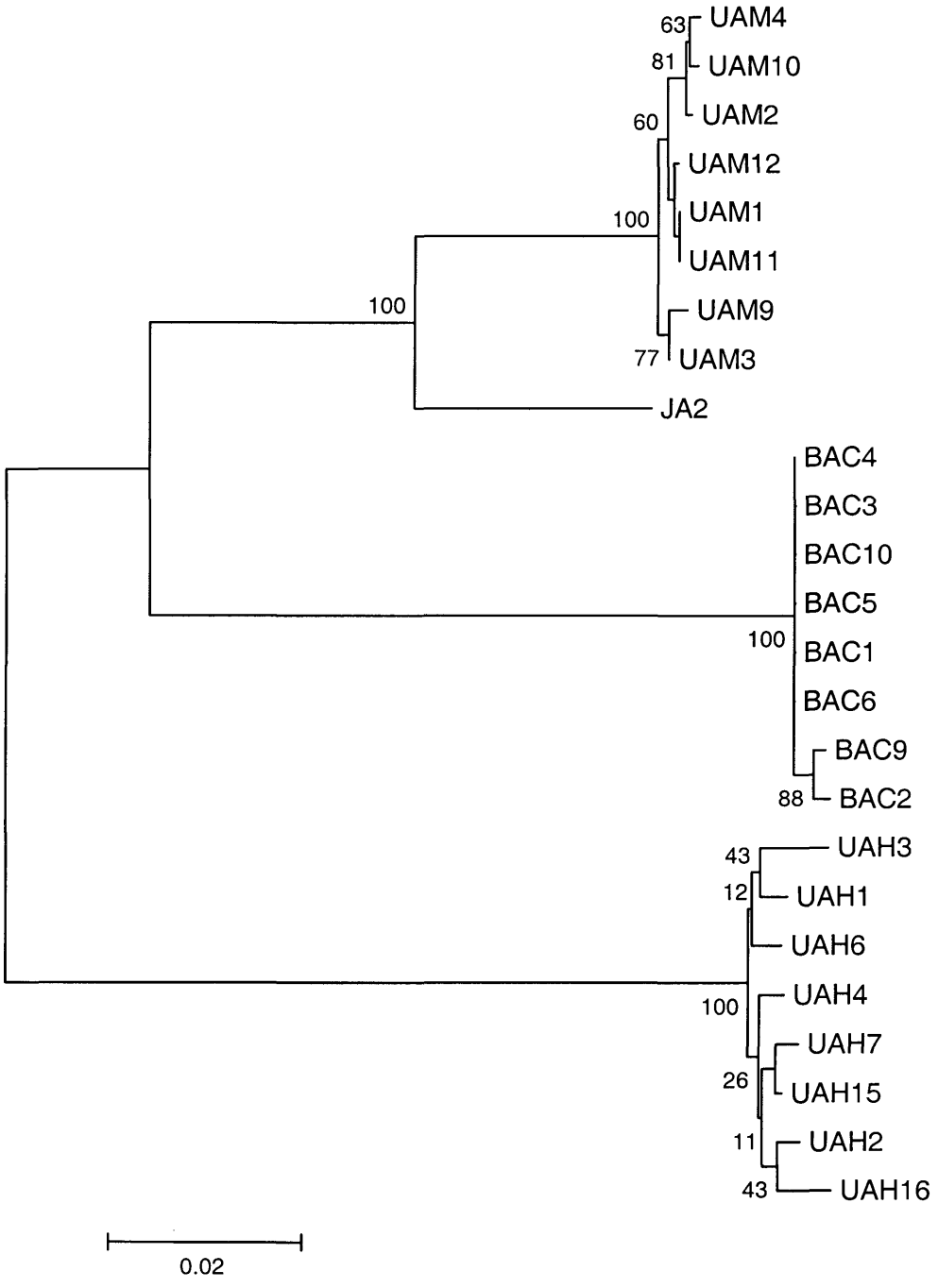

BLG1
62
BLG3 2
83
BLG4
92
ULG8
BLG6
77
98
BLG7
50
ULG10
ULG1
92
59
ULG6
ULG5
BLG5
ULG4
100
BLG8
67
BLG2
30
47
ULG3
0.002

Figure 15. Neighbor-joining tree of control region sequences from *Anchoa* species with bootstrap support values. UAM = *A. mitchilli*; JA = Jamaican *A. hepsetus*; UAH = *A. hepsetus*; BAC = *A. choerostoma*.


UAM4
UAM10
UAM2
UAM12
UAM1
UAM11
UAM9
UAM3
JA2
BAC4
BAC3
BAC10
BAC5
BAC1
BAC6
BAC9
BAC2
UAH3
UAH1
UAH6
UAH4
UAH7
UAH15
UAH2
UAH16
63
81
60
100
100
77
100
88
43
12
100
26
11
43
0.02

collection with 6 to 21 basepair differences between haplotypes (Figure 8B). The neighbor-joining tree of *Haemulon flavolineatum* shows a single, deeply divergent outlying U.S. haplotype (UHF1) (Figure 9). The remaining U.S. and Bermudian haplotypes of this species are all intermingled. The minimum spanning network for *Haemulon sciurus* (Figure 10B) reveals a cluster of Bermudian haplotypes 1 to 9 basepairs removed connected to four U.S. haplotypes. Three of these U.S. haplotypes are between one and 13 basepairs different from the most similar Bermudian haplotype. There are also two deeply divergent, outlying U.S. haplotypes separated from the Bermudian collection by a minimum of 103 basepair changes. For each of the remaining species, the Bermudian and U.S. haplotypes are intermingled with one another in the neighbor-joining trees (Figure 9-14). The neighbor-joining tree of *Anchoa* species (Figure 15) shows that each species comprises an individual clade.

Divergence times based on control region rates of sequence evolution of $3.6 \pm 0.46\%$ ranged from a minimum of 0 years in *Lagodon rhomboides* and *Lutjanus griseus* to a maximum of 277,000 years in *Holocentrus rufus* (Table 9). In addition, a rate of 11.5% was used to calculate a second set of divergence times in *Holocentrus adscensionis*, *H. rufus*, *Lagodon rhomboides*, and *Lutjanus griseus* resulting in a minimum of 0 years in *Lagodon rhomboides* and *Lutjanus griseus* to a maximum of 70,000 years in *Holocentrus rufus*. There appear to be two nodes of mean divergence times among the species. *Haemulon flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus* have mean divergence times between 4,000 and 21,500 years whereas the remaining species (*Haemulon aurolineatum*, *H. sciurus*, *Holocentrus rufus*) have mean divergence times between 56,500 and 189,000 years.

Table 9. Estimated divergence times (in years) using published rates of control region evolution. All ranges include calculations with the standard errors of both the rate and the divergence estimate. For the control region: first times listed are based on a rate of 3.6 ± 0.46%; times in parentheses are based on a rate of 11.5%. Means are listed under each range in bold. For ITS1: first times listed are based on a rate of 1.8 ± 0.23%; times in parentheses on a rate of 5.75%. Means are listed under each range in bold.

Species	Control region	ITS1
Haemulon aurolineatum	65,000 – 168,000 **116,500**	79,000 – 231,000 **155,000**
Haemulon flavolineatum	5,000 – 16,000 **10,500**	13,000 – 27,000 **20,000**
Haemulon sciurus	152,000 – 277,000 **214,500**	99,000 – 188,000 **143,500**
Holocentrus adscensionis	10,000 – 33,000 **21,500** (4,000 – 10,000) **(7,000)**	0 – 18,000 **9,000** (0 – 5000) **(2,500)**
Holocentrus rufus	121,000 – 257,000 **189,000** (43,000 – 70,000) **(56,500)**	59,000 – 202,000 **130,500** (21,000 – 55,000) **(38,000)**
Lagodon rhomboides	0 – 29,000 **14,500** (0 – 8,000) **(4,000)**	0 – 16,000 **8,000** (0 – 4,000) **(2,000)**
Lutjanus griseus	0 – 27,000 **13,500** (0 – 8,000) **(4,000)**	0 – 16,000 **8,000** (0 – 5,000) **(2,500)**

Nuclear DNA

The length of the first internal transcribed spacer (ITS1) region for the species analyzed ranged from 410 bp in the *Anchoa* species to 766 bp in *Lagodon rhomboides* (Table 10). The number of variable positions within ITS1 ranged from 17 in *Lutjanus griseus* to 158 among the *Anchoa* spp. The number of clones sequenced per sample ranged from one to three with a maximum number of three alleles being found per individual (Table 11).

Gene diversity, the equivalent of haplotype diversity for nuclear DNA, ranged from 0.515 in the Bermudian collection of *Haemulon flavolineatum* to 1.00 in several of the species (Table 12). In *Haemulon sciurus*, *Holocentrus adscensionis*, and *Lagodon rhomboides*, both the number of alleles observed and the gene diversity were lower in the Bermudian collection than it was in the U.S. collection. Lower numbers of alleles and gene diversities in the U.S. collections were observed in *Haemulon aurolineatum*, *Haemulon flavolineatum*, *Holocentrus rufus*, and *Lutjanus griseus*. *Anchoa choerostoma*, the endemic Bermudian anchovy, had more alleles (17) and a greater gene diversity (0.971) than *A. hepsetus* (14, 0.881) which has the largest species range of the three *Anchoa* species analyzed. *A. mitchilli*, found only within the United States, had the largest number of alleles observed and the greatest gene diversity (21, 0.986).

Mean nucleotide sequence diversity (π) in Bermudian collections ranged from 0.0013 in *Haemulon flavolineatum* to 0.0144 in *Anchoa choerostoma* (Table 12). Values in U.S. collections ranged from 0.0025 in *Lagodon rhomboides* to 0.0381 in *Anchoa hepsetus* (Table 12). Diversity was greater in the U.S. collection than the Bermudian

Table 10. Length of the first internal transcribed spacer region and the number of variable sites for each species. N=16 for all species except *Lutjanus griseus* (N=15) and the *Anchoa* spp. (N=25).

Species	**Length of ITS1 (bp)**	**Number of variable sites**
Haemulon aurolineatum	676	36
Haemulon flavolineatum	688	40
Haemulon sciurus	691	70
Holocentrus adscensionis	428	20
Holocentrus rufus	456	22
Lagodon rhomboides	766	27
Lutjanus griseus	653	17
Anchoa species	410	158

Table 11. Number of ITS1 alleles observed per individual in relation to the number of clones sequenced per individual.

Species	Location	Number of clones/individual	Number of alleles/individual	Number of individuals
Haemulon aurolineatum	Bermuda	3	1	2
			2	1
			3	5
	U.S.	3	1	2
			2	3
			3	3
Haemulon flavolineatum	Bermuda	3	2	2
			3	2
		2	2	4
	U.S.	3	1	1
		2	1	1
			2	6
Haemulon sciurus	Bermuda	3	3	2
		2	2	6
	U.S.	3	3	3
		2	1	2
			2	5
Holocentrus adscensionis	Bermuda	2	1	2
			2	6
	U.S.	2	1	1
			2	7
Holocentrus rufus	Bermuda	2	2	4
		1	1	4
	U.S.	2	1	1
			2	4
		1	1	3
Lagodon rhomboides	Bermuda	3	1	1
			2	2
			3	5
	U.S.	3	1	1
			2	2
			3	4
		2	2	1
Lutjanus griseus	Bermuda	2	2	8
	U.S.	2	2	7

Anchoa choerostoma	Bermuda	3	2 3	4 4
Anchoa hepsetus	U.S.	3 2	2 3 2	2 5 1
Anchoa hepsetus (J)	Jamaica	3	3	1
Anchoa mitchilli	U.S.	3	2 3	2 6

Table 12. Number of ITS1 alleles, gene diversity, and mean nucleotide sequence diversity (π) for each species by location. Numbers in parentheses are the total number of clones sequenced per location. N=8 for all species in each location with the exception of *Lutjanus griseus* collected in the U.S. (N=7) and *Anchoa hepsetus* collected in Jamaica (N=1).

Species	Location	Number of alleles	Gene diversity	Mean nucleotide sequence diversity (π)
Haemulon aurolineatum	Bermuda	14 (24)	0.884 ± 0.056	0.0077 ± 0.0020
	U.S.	11 (24)	0.837 ± 0.062	0.0040 ± 0.0014
Haemulon flavolineatum	Bermuda	10 (20)	0.711 ± 0.113	0.0013 ± 0.0004
	U.S.	6 (17)	0.515 ± 0.145	0.0097 ± 0.0017
Haemulon sciurus	Bermuda	10 (17)	0.868 ± 0.070	0.0040 ± 0.0012
	U.S.	12 (17)	0.941 ± 0.043	0.0266 ± 0.0039
Holocentrus adscensionis	Bermuda	11 (16)	0.942 ± 0.041	0.0068 ± 0.0023
	U.S.	13 (16)	0.975 ± 0.029	0.0079 ± 0.0022
Holocentrus rufus	Bermuda	11 (12)	0.985 ± 0.040	0.0109 ± 0.0028
	U.S.	8 (13)	0.897 ± 0.067	0.0044 ± 0.0016
Lagodon rhomboides	Bermuda	10 (24)	0.822 ± 0.061	0.0024 ± 0.0009
	U.S.	13 (23)	0.846 ± 0.071	0.0025 ± 0.0007
Lutjanus griseus	Bermuda	7 (16)	0.792 ± 0.089	0.0025 ± 0.0009
	U.S.	7 (14)	0.758 ± 0.116	0.0028 ± 0.0009
Anchoa species	*A. choerostoma*	17 (24)	0.971 ± 0.019	0.0144 ± 0.0036
	A. hepsetus	14 (23)	0.881 ± 0.062	0.0381 ± 0.0058
	A. hepsetus (J)	3 (3)	1.00 ± 0.272	0.0400 ± 0.0083
	A. mitchilli	21 (24)	0.986 ± 0.018	0.0254 ± 0.0052

collection for *Haemulon flavolineatum*, *Haemulon sciurus* (an order of magnitude larger), *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus*. The opposite pattern was observed for *Haemulon aurolineatum* and *Holocentrus rufus* (values were an order of magnitude larger in Bermuda). *Anchoa choerostoma*, the endemic Bermudian anchovy, had the lowest diversity of the anchovy species (0.0144). The value was largest in the Jamaican *A. hepsetus* (0.04). However, this is most likely an artifact of the value the Jamaican *A. hepsetus* being calculated from three clones of a single individual.

Similar to results obtained with the mitochondrial control region, estimates of divergence between populations using ITS1 varied by an order of magnitude. The mean nucleotide sequence divergence between Bermudian and U.S. collections ranged from 0.0025 in *Lagodon rhomboides* to 0.0177 in *Haemulon sciurus* (Table 13). The net mean nucleotide sequence divergence varied from 0.000091 in *Holocentrus adscensionis* to 0.002612 in *Haemulon aurolineatum* (Table 14). For *Holocentrus adscensionis* and *Lutjanus griseus*, the standard error is an order of magnitude larger than the actual value of divergence suggesting that there is no detectable signal between Bermudian and U.S. populations of these species.

Mean nucleotide sequence divergence among the *Anchoa* species ranged from 0.052 between *A. hepsetus* (Jamaica) and *A. mitchilli* to 0.215 between *A. choerostoma* and *A. hepsetus* (U.S.) (Table 14). Net mean nucleotide sequence divergence ranged from 0.019 between *A. hepsetus* (Jamaica) and *A. mitchilli* to 0.188 between *A. choerostoma* and *A. hepsetus* (U.S.). The endemic Bermudian anchovy, *A. choerostoma*, has its lowest mean nucleotide sequence divergence (0.146) and lowest net mean nucleotide sequence divergence (0.126) with *A. mitchilli*.

Table 13. Mean nucleotide sequence divergence and net mean nucleotide sequence divergence between Bermudian and U.S. populations of each species for the ITS1 region.

Species	Mean nucleotide sequence divergence	Net mean nucleotide sequence divergence
Haemulon aurolineatum	0.0085 ± 0.0020	0.002612 ± 0.002014
Haemulon flavolineatum	0.0059 ± 0.0009	0.000344 ± 0.000081
Haemulon sciurus	0.0177 ± 0.0024	0.002483 ± 0.000469
Holocentrus adscensionis	0.0074 ± 0.0022	0.000091 ± 0.000199
Holocentrus rufus	0.0098 ± 0.0027	0.002180 ± 0.000993
Lagodon rhomboides	0.0025 ± 0.0008	0.000122 ± 0.000122
Lutjanus griseus	0.0027 ± 0.0007	0.000097 ± 0.000161

Table 14. Mean nucleotide sequence divergence (below the diagonal) and net mean nucleotide sequence divergence (above the diagonal) between the species of *Anchoa* for the ITS1 region.

	A. mitchilli	*A. hepsetus*	*A. hepsetus (Jamaica)*	*A. choerostoma*
A. mitchilli	--	0.172 ± 0.023	0.019 ± 0.005	0.126 ± 0.020
A. hepsetus	0.204 ± 0.024	--	0.175 ± 0.024	0.188 ± 0.026
A. hepsetus (Jamaica)	0.052 ± 0.009	0.214 ± 0.025	--	0.129 ± 0.021
A. choerostoma	0.146 ± 0.022	0.215 ± 0.028	0.156 ± 0.022	--

Fst values ranged from -0.00090 (0) in *Holocentrus adscensionis* to 0.81798 among the *Anchoa* spp. (Table 15). Fst values were significant between Bermudian and U.S. populations of *Haemulon aurolineatum* (0.26496, $p<0.01$), *Haemulon flavolineatum* (0.06154, $p=0.031$), *Haemulon sciurus* (0.13137, $p=0.031$), *Holocentrus rufus* (0.16101, $p=0.001$), and *Lutjanus griseus* (0.23078, $p<0.01$). For these five species, AMOVA revealed that 26.5%, 6.15%, 13.14%, 16.1%, and 23.08% of the variance was found among populations, respectively. The value of Fst between the *Anchoa* species was also significant (0.81798, $p<0.01$), with 81.8% of the variance found among the species. Other comparisons were not significant.

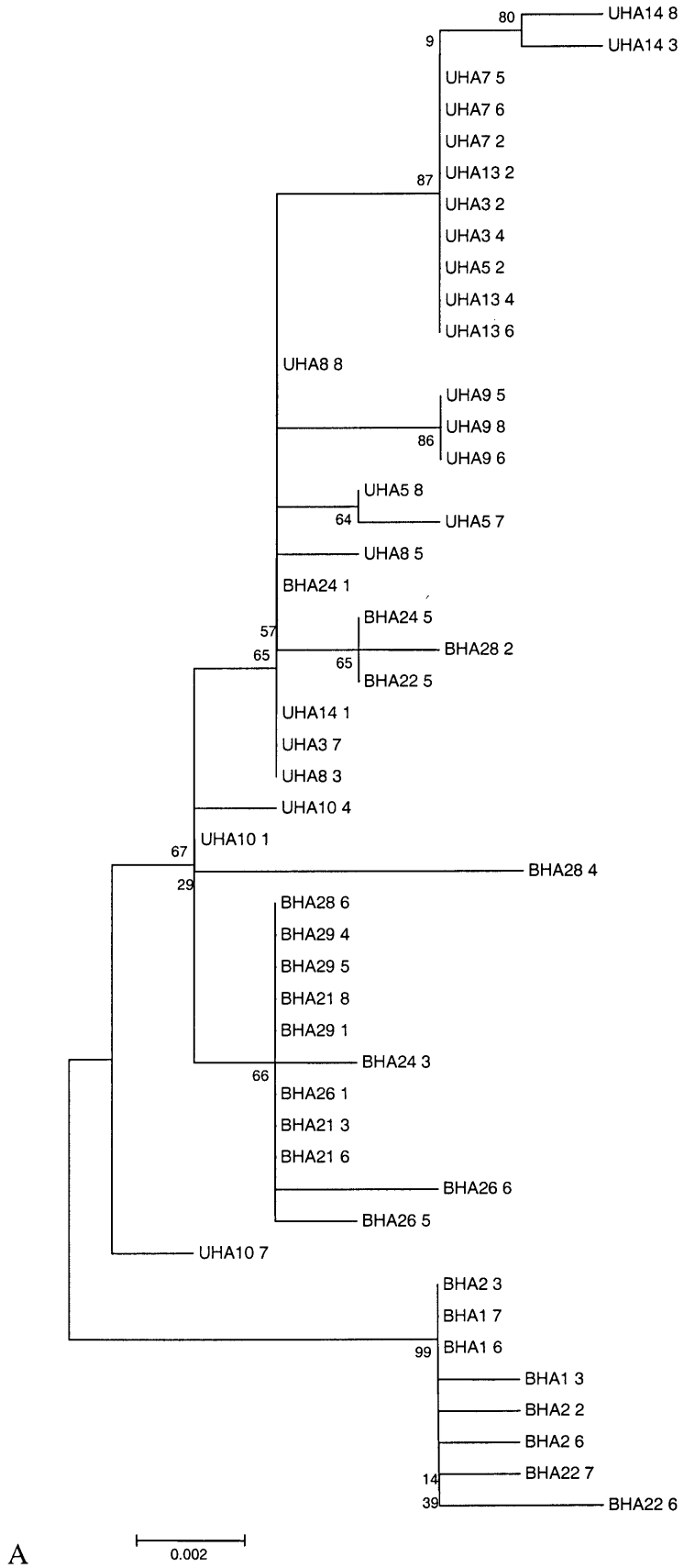
Neighbor-joining trees based on ITS1 sequence data for each species and minimum spanning networks for those species with significant Fst values are shown in Figures 16 – 23. For all of the species examined, clones from the same individual are in many cases intermingled with clones from other individuals.

There are two large clusters of Bermudian *Haemulon aurolineatum* alleles that comprise a central, shared allele with the other alleles only one or two basepairs different (Figure 16B). One of these clusters is revealed to be its own separate clade in the neighbor-joining tree (Figure 16A), although a single allele (BHA22_5) from individual BHA22 in this clade is found intermingled with the remaining alleles. The second Bermudian cluster and five remaining Bermudian alleles all stem from two very similar U.S. alleles (one basepair different). In the neighbor-joining tree and minimum spanning network of *Haemulon flavolineatum*, alleles from the U.S. individual UHF1 are deeply divergent outliers in a single, separate clade, while all other Bermudian and U.S. alleles are intermingled with one another (Figure 17A,B). This is the same individual that was

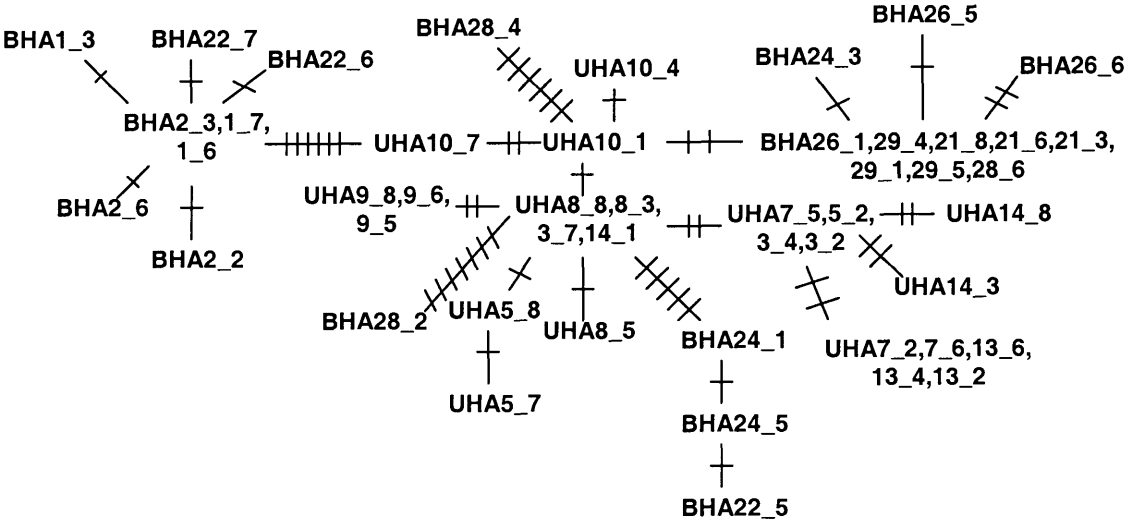
Table 15. Fst and AMOVA results based upon ITS1 sequences for each species.

Species	**Fst**	**AMOVA variance among populations**	**AMOVA variance within populations**
Haemulon aurolineatum	0.26496 (p=0.000)	26.5%	73.5%
Haemulon flavolineatum	0.06154 (p=0.031)	6.15%	93.85%
Haemulon sciurus	0.13137 (p=0.031)	13.14%	86.86%
Holocentrus adscensionis	-0.00090 (p=0.413)	-0.09%	100.09%
Holocentrus rufus	0.16101 (p=0.001)	16.1%	83.9%
Lagodon rhomboides	0.03307 (p=0.054)	3.31%	96.69%
Lutjanus griseus	0.23078 (p=0.000)	23.08%	76.92%
Anchoa spp.	0.81798 (p=0.000)	81.8%	18.2%

Figure 16.A) Neighbor-joining tree of ITS1 sequences from *Haemulon aurolineatum* with bootstrap support values. B) Minimum spanning network of Bermudian and U.S. haplotypes. BHA = Bermudian samples; UHA = U.S. samples.

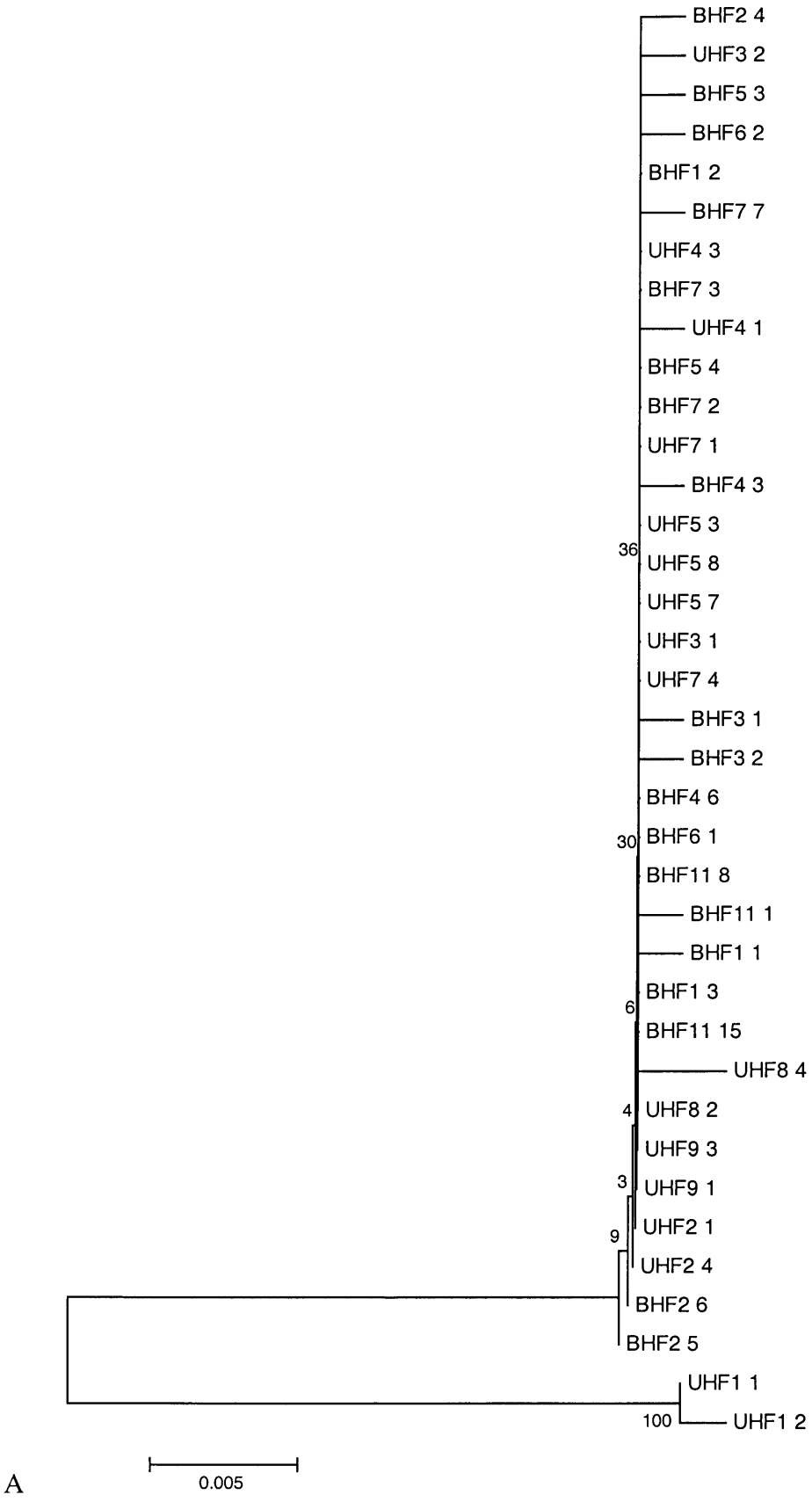

80
UHA14 8
9
UHA14 3
UHA7 5
UHA7 6
UHA7 2
87
UHA13 2
UHA3 2
UHA3 4
UHA5 2
UHA13 4
UHA13 6
UHA8 8
UHA9 5
UHA9 8
86
UHA9 6
UHA5 8
64
UHA5 7
UHA8 5
BHA24 1
BHA24 5
57
65
BHA28 2
65
BHA22 5
UHA14 1
UHA3 7
UHA8 3
UHA10 4
UHA10 1
67
BHA28 4
29
BHA28 6
BHA29 4
BHA29 5
BHA21 8
BHA29 1
BHA24 3
66
BHA26 1
BHA21 3
BHA21 6
BHA26 6
BHA26 5
UHA10 7
BHA2 3
BHA1 7
99
BHA1 6
BHA1 3
BHA2 2
BHA2 6
14
BHA22 7
39
BHA22 6
0.002

A

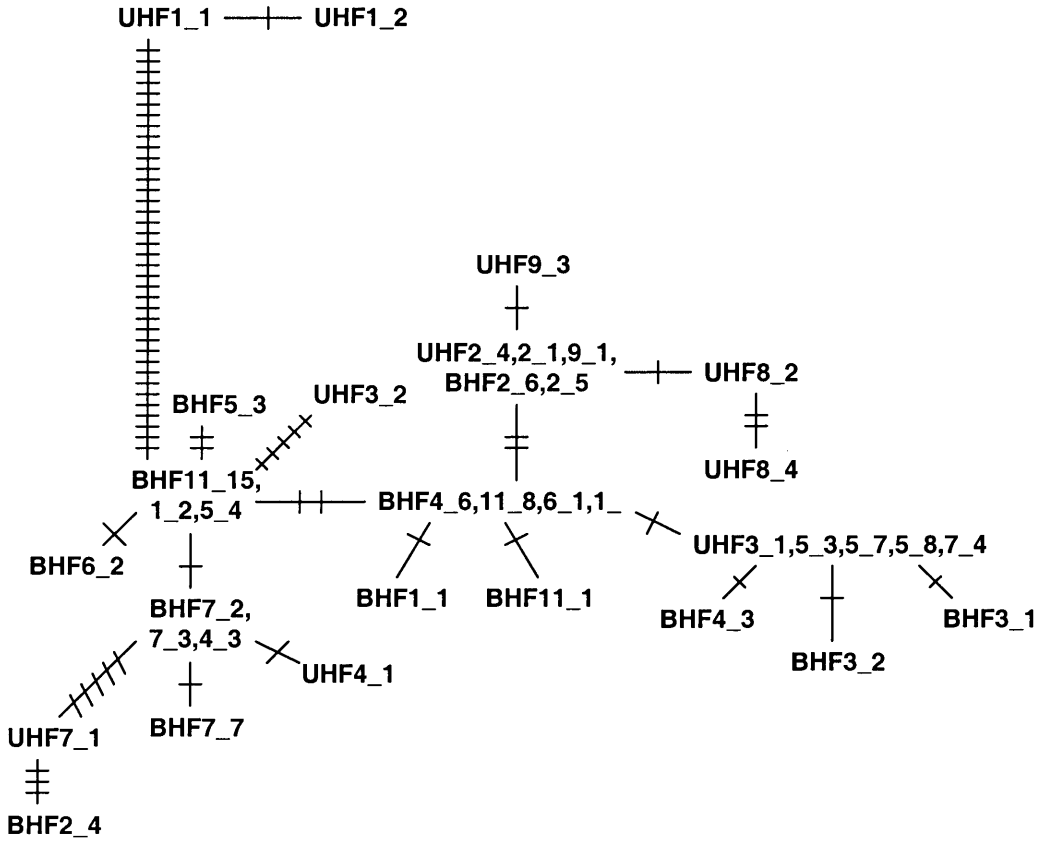

BHA1_3
BHA22_7
BHA22_6
BHA2_3,1_7,
1_6
BHA2_6
BHA2_2
BHA28_4
UHA10_4
UHA10_7
UHA10_1
UHA9_8,9_6,
9_5
UHA8_8,8_3,
3_7,14_1
BHA28_2
UHA5_8
UHA8_5
UHA5_7
BHA24_1
BHA24_5
BHA22_5
BHA26_5
BHA24_3
BHA26_6
BHA26_1,29_4,21_8,21_6,21_3,
29_1,29_5,28_6
UHA7_5,5_2,
3_4,3_2
UHA14_8
UHA14_3
UHA7_2,7_6,13_6,
13_4,13_2

B

Figure 17. A) Neighbor-joining tree of ITS1 sequences from *Haemulon flavolineatum* with bootstrap support values. B) Minimum spanning network of Bermudian and U.S. haplotypes. BHF = Bermudian samples; UHF = U.S. samples.

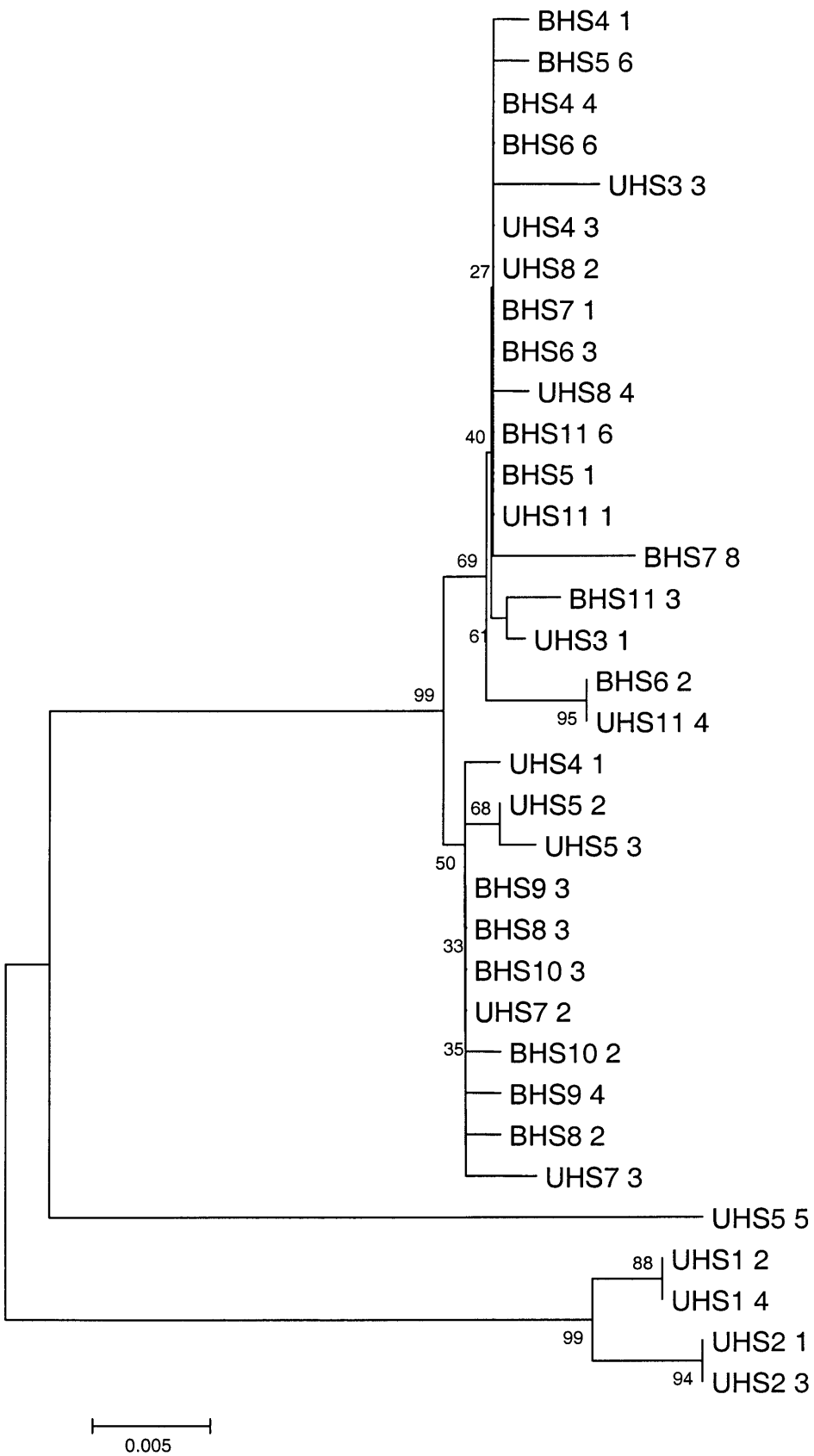

BHF2 4
UHF3 2
BHF5 3
BHF6 2
BHF1 2
BHF7 7
UHF4 3
BHF7 3
UHF4 1
BHF5 4
BHF7 2
UHF7 1
BHF4 3
UHF5 3
36
UHF5 8
UHF5 7
UHF3 1
UHF7 4
BHF3 1
BHF3 2
BHF4 6
30
BHF6 1
BHF11 8
BHF11 1
BHF1 1
BHF1 3
6
BHF11 15
UHF8 4
4
UHF8 2
UHF9 3
3
UHF9 1
9
UHF2 1
UHF2 4
BHF2 6
BHF2 5
UHF1 1
100
UHF1 2
0.005

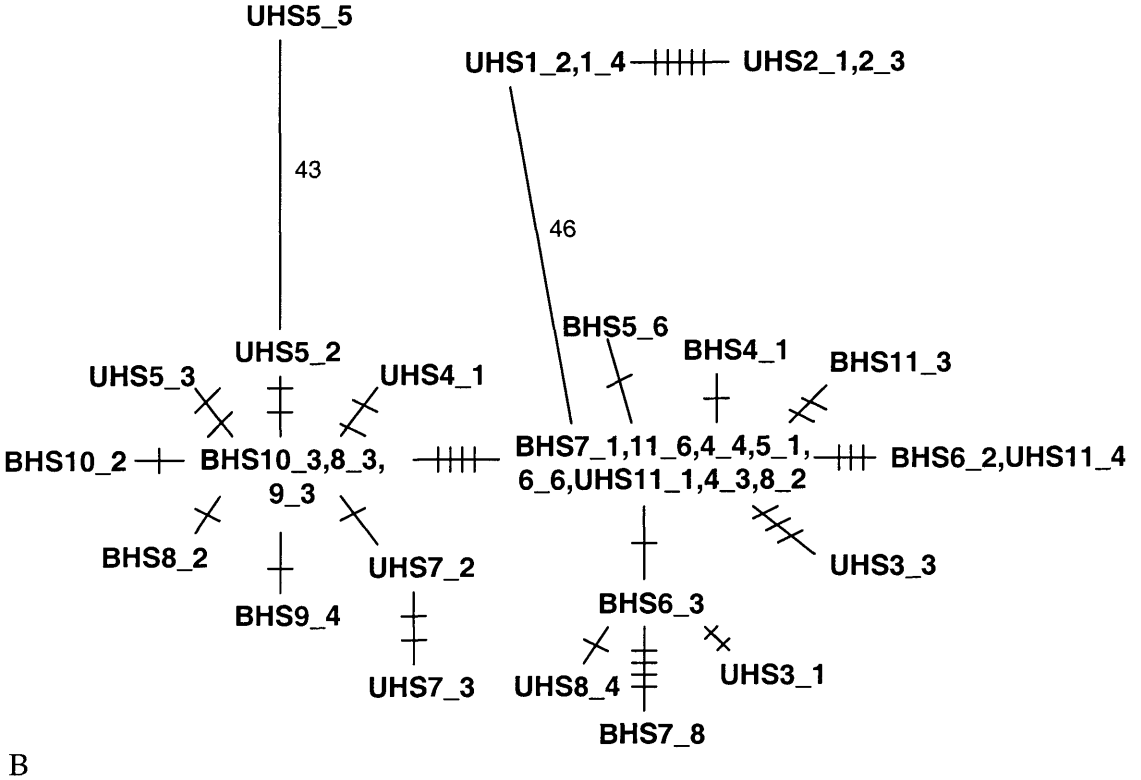
A


UHF1_1
UHF1_2
UHF9_3
UHF2_4,2_1,9_1,
BHF2_6,2_5
UHF8_2
BHF5_3
UHF3_2
UHF8_4
BHF11_15,
1_2,5_4
BHF4_6,11_8,6_1,1_
UHF3_1,5_3,5_7,5_8,7_4
BHF6_2
BHF1_1
BHF11_1
BHF4_3
BHF3_1
BHF7_2,
7_3,4_3
BHF3_2
UHF4_1
UHF7_1
BHF7_7
BHF2_4

B

Figure 18. A) Neighbor-joining tree of ITS1 sequences from *Haemulon sciurus* with bootstrap support values. B) Minimum spanning network of Bermudian and U.S. haplotypes. BHS = Bermudian samples; UHS = U.S. samples.


BHS4 1
BHS5 6
BHS4 4
BHS6 6
UHS3 3
UHS4 3
27
UHS8 2
BHS7 1
BHS6 3
UHS8 4
40
BHS11 6
BHS5 1
UHS11 1
69
BHS7 8
BHS11 3
61
UHS3 1
99
BHS6 2
95
UHS11 4
UHS4 1
68
UHS5 2
UHS5 3
50
BHS9 3
BHS8 3
33
BHS10 3
UHS7 2
35
BHS10 2
BHS9 4
BHS8 2
UHS7 3
UHS5 5
88
UHS1 2
UHS1 4
99
UHS2 1
94
UHS2 3
A
0.005


UHS5_5
UHS1_2,1_4
UHS2_1,2_3
43
46
BHS5_6
BHS4_1
BHS11_3
UHS5_2
UHS5_3
UHS4_1
BHS10_2
BHS10_3,8_3,
9_3
BHS7_1,11_6,4_4,5_1,
6_6,UHS11_1,4_3,8_2
BHS6_2,UHS11_4
BHS8_2
UHS7_2
UHS3_3
BHS9_4
BHS6_3
UHS7_3
UHS8_4
UHS3_1
BHS7_8

B

Figure 19. Neighbor-joining tree of ITS1 sequences from *Holocentrus adscensionis* with bootstrap support values. BHOA = Bermudian samples; UHOA = U.S. samples.

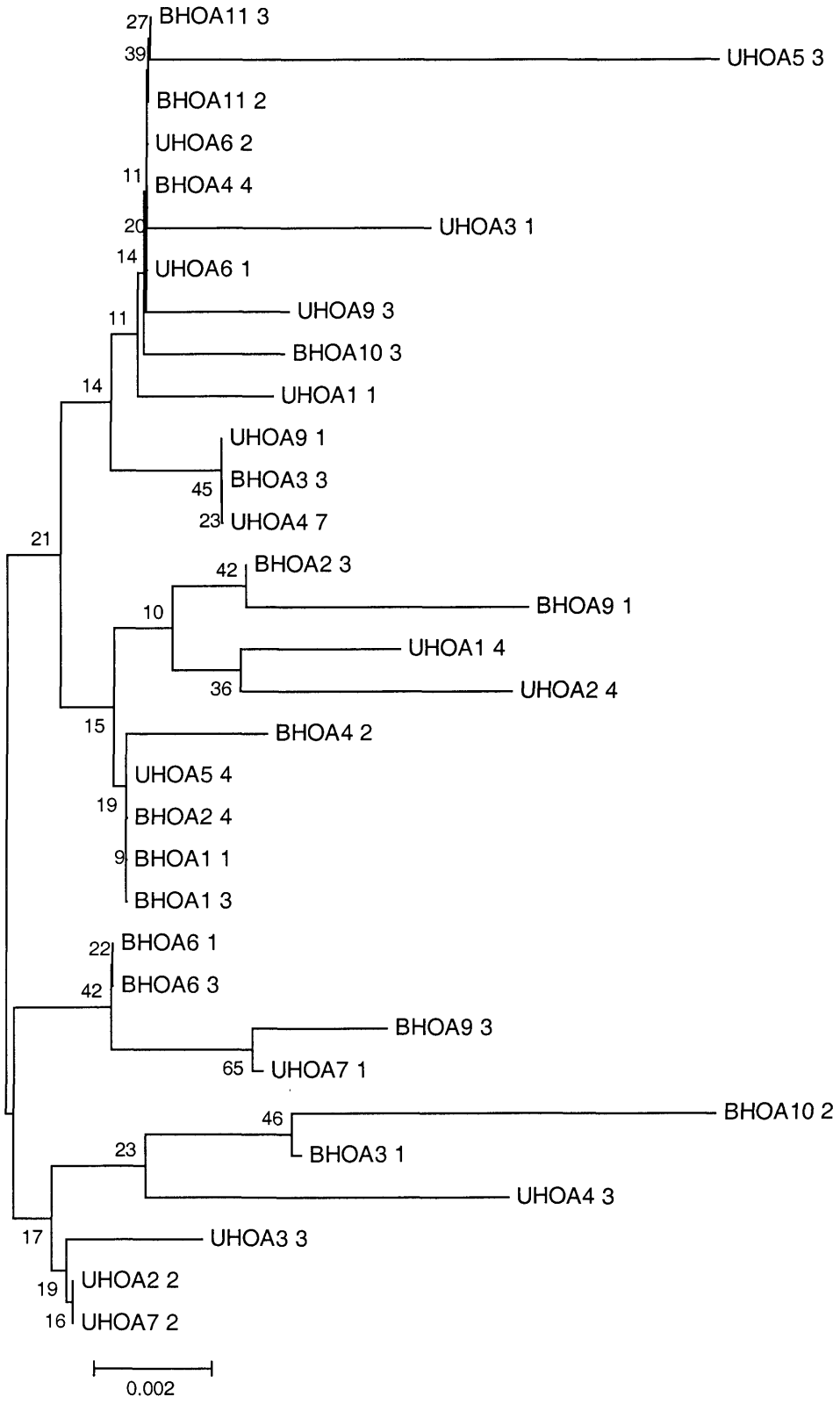
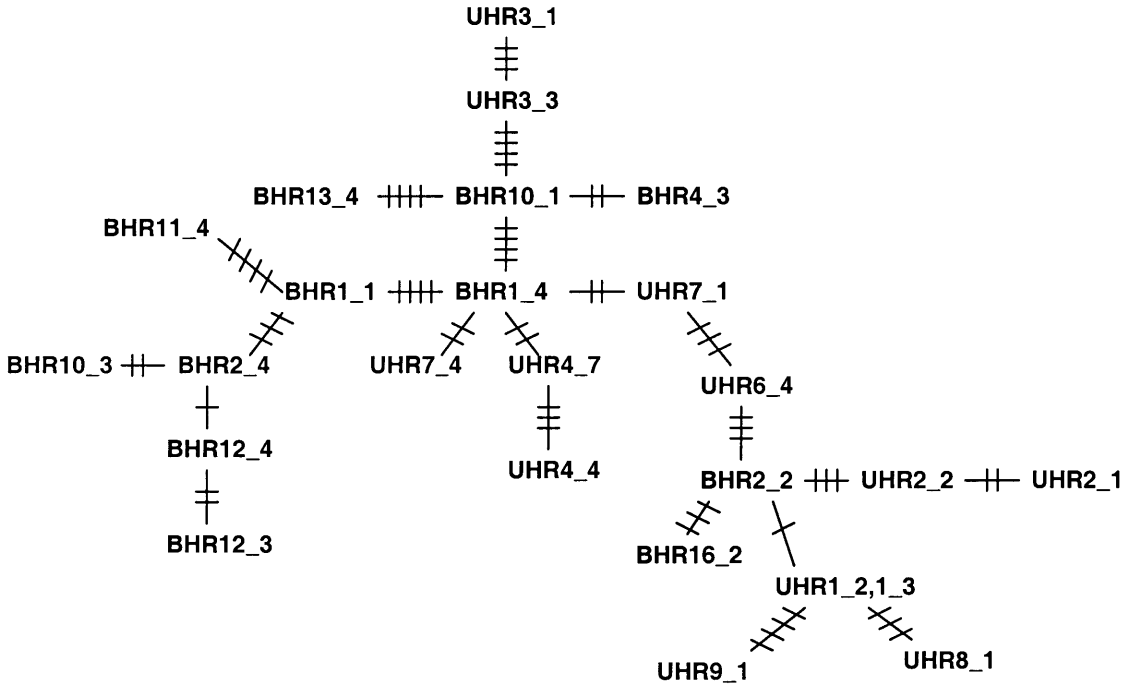

BHOA11 3
UHOA5 3
BHOA11 2
UHOA6 2
BHOA4 4
UHOA3 1
UHOA6 1
UHOA9 3
BHOA10 3
UHOA1 1
UHOA9 1
BHOA3 3
UHOA4 7
BHOA2 3
BHOA9 1
UHOA1 4
UHOA2 4
BHOA4 2
UHOA5 4
BHOA2 4
BHOA1 1
BHOA1 3
BHOA6 1
BHOA6 3
BHOA9 3
UHOA7 1
BHOA10 2
BHOA3 1
UHOA4 3
UHOA3 3
UHOA2 2
UHOA7 2
0.002

Figure 20. A) Neighbor-joining tree of ITS1 sequences from *Holocentrus rufus* with bootstrap support values. B) Minimum spanning network of Bermudian and U.S. haplotypes. BHR = Bermudian samples; UHR = U.S. samples.


31
UHR1 3
UHR8 1
9
UHR1 2
BHR16 2
27
BHR2 2
17
UHR9 1
UHR6 4
44
UHR3 1
65
UHR3 3
63
BHR13 4
65
UHR2 1
73
UHR2 2
BHR10 1
BHR1 4
47
BHR4 3
UHR4 7
45
27
UHR7 4
17
UHR4 4
40
UHR7 1
41
BHR11 4
BHR1 1
87
BHR2 4
48
BHR10 3
40
BHR12 3
50
BHR12 4
0.002

A


UHR3_1
UHR3_3
BHR13_4
BHR10_1
BHR4_3
BHR11_4
BHR1_1
BHR1_4
UHR7_1
BHR10_3
BHR2_4
UHR7_4
UHR4_7
UHR6_4
BHR12_4
UHR4_4
BHR2_2
UHR2_2
UHR2_1
BHR16_2
BHR12_3
UHR1_2,1_3
UHR9_1
UHR8_1

B

Figure 21. Neighbor-joining tree of ITS1 sequences from *Lagodon rhomboides* with bootstrap support values. BLR = Bermudian samples; ULR = U.S. samples.

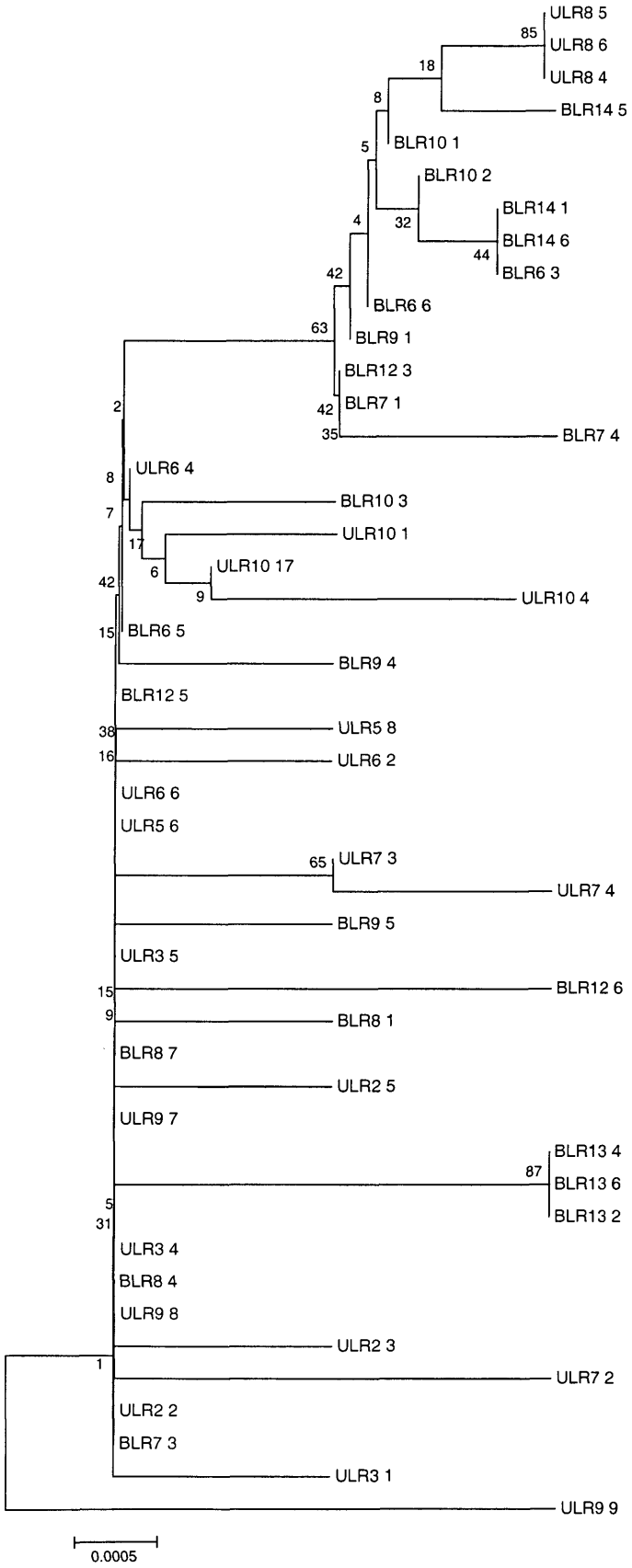
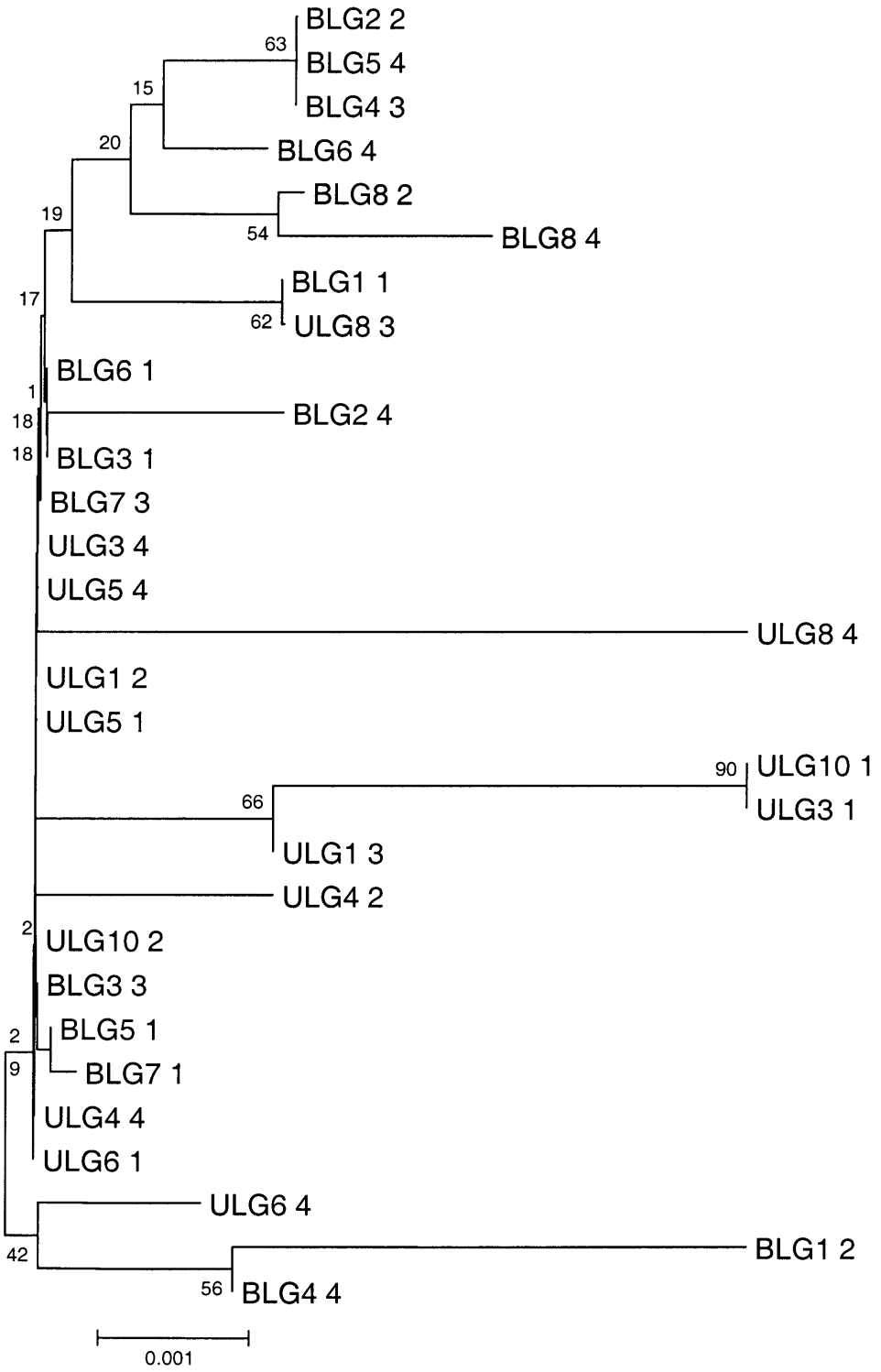
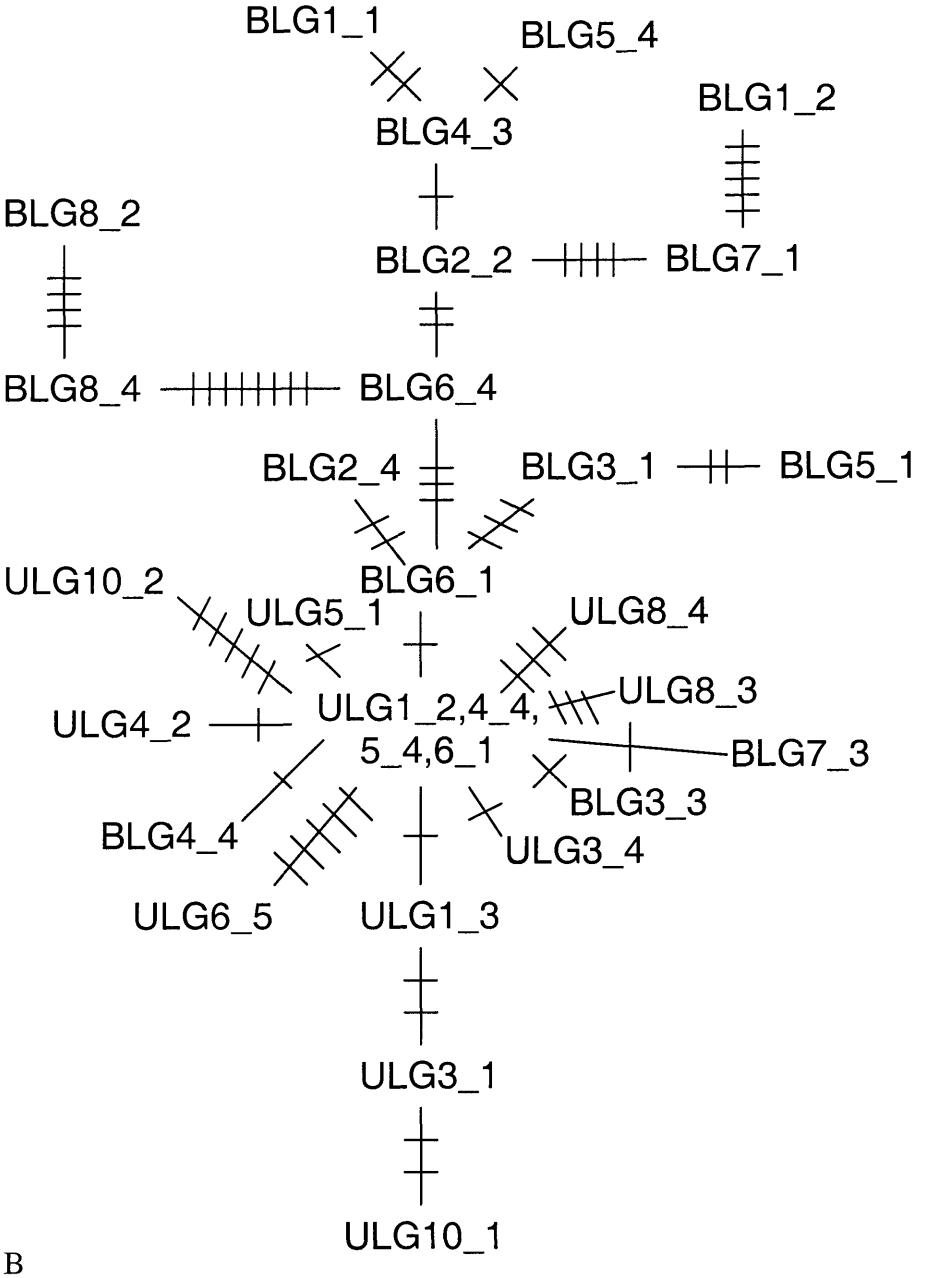

ULR8 5
85
ULR8 6
18
ULR8 4
8
BLR14 5
5
BLR10 1
BLR10 2
4
32
BLR14 1
BLR14 6
44
BLR6 3
42
BLR6 6
63
BLR9 1
BLR12 3
2
42
BLR7 1
35
BLR7 4
8
ULR6 4
7
BLR10 3
17
ULR10 1
6
ULR10 17
42
9
ULR10 4
15
BLR6 5
BLR9 4
BLR12 5
38
ULR5 8
16
ULR6 2
ULR6 6
ULR5 6
65
ULR7 3
ULR7 4
BLR9 5
ULR3 5
15
BLR12 6
9
BLR8 1
BLR8 7
ULR2 5
ULR9 7
BLR13 4
87
BLR13 6
5
BLR13 2
31
ULR3 4
BLR8 4
ULR9 8
ULR2 3
1
ULR7 2
ULR2 2
BLR7 3
ULR3 1
ULR9 9
0.0005

Figure 22. A) Neighbor-joining tree of ITS1 sequences from *Lutjanus griseus* with bootstrap support values. B) Minimum spanning network of Bermudian and U.S. haplotypes. BLG = Bermudian samples; ULG = U.S. samples.

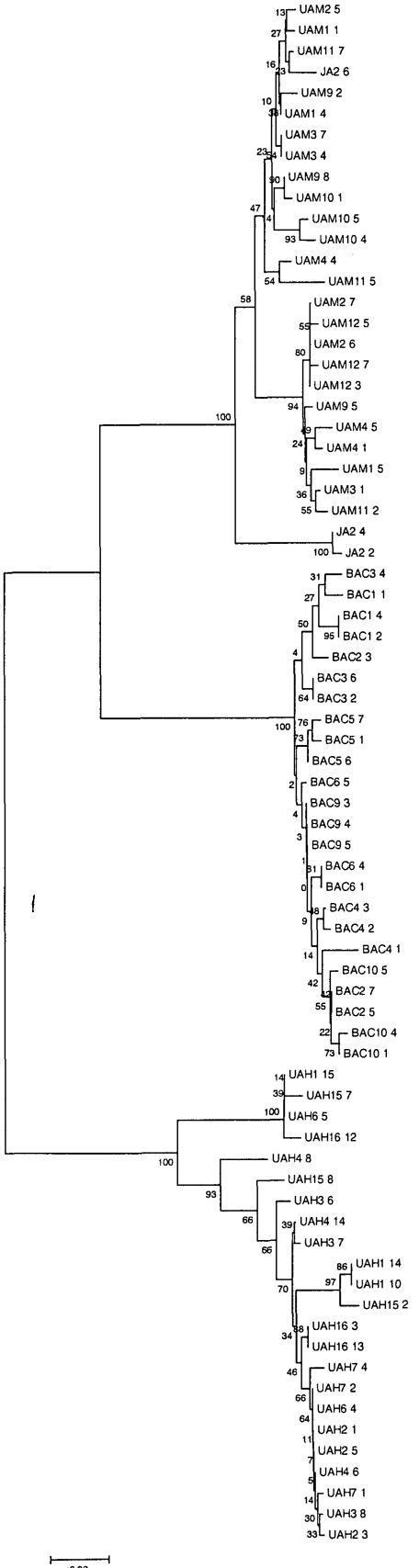

BLG2 2
63
BLG5 4
15
BLG4 3
20
BLG6 4
BLG8 2
19
54
BLG8 4
BLG1 1
17
62
ULG8 3
BLG6 1
1
BLG2 4
18
18
BLG3 1
BLG7 3
ULG3 4
ULG5 4
ULG8 4
ULG1 2
ULG5 1
90
ULG10 1
66
ULG3 1
ULG1 3
ULG4 2
2
ULG10 2
BLG3 3
2
BLG5 1
9
BLG7 1
ULG4 4
ULG6 1
ULG6 4
42
BLG1 2
56
BLG4 4
0.001

A


BLG1_1
BLG5_4
BLG4_3
BLG1_2
BLG8_2
BLG2_2
BLG7_1
BLG8_4
BLG6_4
BLG2_4
BLG3_1
BLG5_1
BLG6_1
ULG10_2
ULG5_1
ULG8_4
ULG8_3
ULG4_2
ULG1_2,4_4,
5_4,6_1
BLG7_3
BLG3_3
BLG4_4
ULG3_4
ULG6_5
ULG1_3
ULG3_1
ULG10_1

B

Figure 23. Neighbor-joining tree of ITS1 sequences from *Anchoa* species with bootstrap support values. UAM = *A. mitchilli*; JA = Jamaican *A. hepsetus*; UAH = *A. hepsetus*; BAC = *A. choerostoma*.


UAM2 5
UAM1 1
UAM11 7
JA2 6
UAM9 2
UAM1 4
UAM3 7
UAM3 4
UAM9 8
UAM10 1
UAM10 5
UAM10 4
UAM4 4
UAM11 5
UAM2 7
UAM12 5
UAM2 6
UAM12 7
UAM12 3
UAM9 5
UAM4 5
UAM4 1
UAM1 5
UAM3 1
UAM11 2
JA2 4
JA2 2
BAC3 4
BAC1 1
BAC1 4
BAC1 2
BAC2 3
BAC3 6
BAC3 2
BAC5 7
BAC5 1
BAC5 6
BAC6 5
BAC9 3
BAC9 4
BAC9 5
BAC6 4
BAC6 1
BAC4 3
BAC4 2
BAC4 1
BAC10 5
BAC2 7
BAC2 5
BAC10 4
BAC10 1
UAH1 15
UAH15 7
UAH6 5
UAH16 12
UAH4 8
UAH15 8
UAH3 6
UAH4 14
UAH3 7
UAH1 14
UAH1 10
UAH15 2
UAH16 3
UAH16 13
UAH7 4
UAH7 2
UAH6 4
UAH2 1
UAH2 5
UAH4 6
UAH7 1
UAH3 8
UAH2 3
0.02

considered an outlier using mitochondrial control region data (Figure 10). The alleles of two U.S. individuals of *Haemulon sciurus* (UHS1,2) are divergent outliers, separated from the remaining Bermudian and U.S. alleles by 46 basepair changes (Figures 18A,B). Similar to that observed in *Haemulon flavolineatum*, these are the same individuals that were considered outliers using mitochondrial control region data (Figures 1A,B). In addition, a single allele of the U.S. individual UHS5 is also very divergent from the remaining *Haemulon sciurus* alleles, separated by 43 basepair changes, while the other allele from this individual is only separated by two basepair changes (Figure 18B).

All alleles from both Bermudian and U.S. individuals of the squirrelfish *Holocentrus adscensionis* are intermingled (Figure 19), while there is a clade of very similar Bermudian alleles from individuals of the other squirrelfish examined, *Holocentrus rufus*, that is separate from all other alleles (Figures 20A,B). Two of these individuals also have alleles that are intermingled within the rest of the alleles observed.

The neighbor-joining tree of *Lagodon rhomboides* reveals a small clade of Bermudian alleles that also contains all alleles from the U.S. individual ULR8. The remaining Bermudian and U.S. alleles are intermingled with the exception of a single, outlying allele from the U.S. individual ULR9 (Figure 21). For *Lutjanus griseus*, a large cluster of Bermudian alleles and three single Bermudian alleles are connected to one U.S. allele, each by a difference of only one basepair (Figures 22A,B).

The neighbor-joining tree of *Anchoa* species (Figure 23) shows that each species comprises an individual clade.

Divergence times for ITS1 based on rates of sequence evolution of

1.8 ± 0.23% ranged from a minimum of 0 years in *Holocentrus adscensionis*, *Lagodon rhomboides* and *Lutjanus griseus* to a maximum of 231,000 years in *Haemulon aurolineatum* (Table 9). A rate of 5.75% was used to calculate a second set of divergence times in *Holocentrus adscensionis*, *H. rufus*, *Lagodon rhomboides*, and *Lutjanus griseus* resulting in a minimum of 0 years in *Holocentrus adscensionis*, *Lagodon rhomboides* and *Lutjanus griseus* to a maximum of 55,000 years in *Holocentrus rufus*. As was observed in divergence times based on the mitochondrial control region, there appear to be two nodes of mean divergence times among the species. *Haemulon flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus* have mean divergence times between 2,000 and 20,000 years whereas the remaining species (*Haemulon aurolineatum*, *H. sciurus*, *Holocentrus rufus*) have mean divergence times between 38,000 and 155,000 years. Divergence times estimated from ITS1 sequences were not significantly different from those estimated using control region sequences ($F=0.16$, $p=0.694$).

Discussion

Diversity within Bermuda

The first objective of this study was to determine if Bermudian populations of tropical shorefishes represent recent colonists that arrived at the end of the last ice age approximately 18,000 years ago. Sequence analysis of the mitochondrial control region revealed considerably high gene diversity within Bermudian collections (0.857 to 1.00), with the exception of the endemic Bermudian anchovy, *Anchoa choerostoma* (0.464). Differences among haplotypes were not great, however, as mean nucleotide sequence diversities were between 0.0015 and 0.0401. These results are quite similar to those obtained by Muss et al. (2001) in populations of *Ophioblennius* throughout the Atlantic where haplotype diversity was "very high" between 0.803 and 1.00, with the exception of *O. atlanticus* in the Azores (0.417), and "low to moderate" mean nucleotide sequence diversities between 0.0011 and 0.0122. Analysis of the nuclear ITS1 region resulted in the same pattern with high gene diversities (0.711 to 0.985) and low to moderate mean nucleotide sequence diversities (0.0013 to 0.0144) in all Bermudian collections.

This combination of high haplotype diversity with low nucleotide sequence diversity is not the pattern most closely associated with recent recolonization as described by Grant and Bowen (1998). For such populations, one would expect to find both low haplotype diversity and low mean nucleotide sequence diversity resulting from an original founder population and not enough time having passed for the accumulation of large amounts of variation. *Ophioblennius atlanticus* in the Azores, a location where it is believed warm-water fauna were extirpated during the late Pleistocene glaciation, demonstrate this pattern with a low haplotype diversity ($h = 0.417$) and a low nucleotide

sequence diversity ($\pi = 0.0011$) based on cytochrome *b* sequence data (Muss et al. 2001). According to the authors, the shallow diversity is a vestige of recent colonization from elsewhere in the Lusitanian biogeographic province (Figure 4).

Even though the mitochondrial control region evolves at a faster rate than cytochrome *b*, it would be expected that Bermudian populations would exhibit the same general pattern of low haplotype and mean nucleotide sequence diversities if they were the result of recent post-glacial colonization. However, the combination of high haplotype diversity and low nucleotide sequence diversity does not preclude the possibility of recent recolonization. It may be indicative of rapid population growth following a founder event or a period of low effective population size (Grant and Bowen 1998). This rapid population growth can result in a large population with many haplotypes separated by only 1 or 2 basepair changes. Indeed, the minimum spanning network of *Haemulon aurolineatum* samples (Figure 8B) shows this pattern. The Bermudian *H. aurolineatum* haplotypes cluster together and are all zero to two mutations removed from one another. Bermudian haplotypes are connected to a single U.S. haplotype by one mutation. The U.S. samples themselves are more structured than those of Bermuda and are separated by 6 to 21 mutations. The pattern in the U.S. samples is indicative of a large, stable population. Although this pattern is not observed in the minimum spanning networks of all species examined, the overall evidence within mitochondrial DNA of Bermudian samples suggests rapid population growth following a founder event.

A similar pattern of high gene diversity and low nucleotide sequence diversity is found in the analysis of ITS1. This region is part of the multi-copy ribosomal RNA gene

complex; as such, there may be hundreds of copies present within an individual genome. It is therefore not unexpected to find moderate gene diversity within each species as more than two alleles may be found within an individual (Palumbi and Baker 1994). Through the process of concerted evolution, one would also expect to find relatively low nucleotide sequence diversities within each species. However, recent studies of variation in both ITS1 and ITS2 have demonstrated high intraspecific variability. Harris and Crandall (2000) sequenced between four and ten clones from each of nine individuals of *Orconectes* (crayfish) and found greater pairwise distances within individuals (0-0.00503) and among individuals of the same population (0.00443) than among populations within a given species (0.00069-0.00435). Vogler and DeSalle (1994) sequenced three to five clones from each of 12 individuals of *Cincindela dorsalis* (tiger beetle). Forty-two of the 50 clones differed by at least one nucleotide although the geographic partioning of ITS1 alleles was congruent with that of previous mitochondrial DNA analyses. In this study, between 1 and 3 alleles were found within each individual with the maximum number of clones sequenced from any individual being three.

Similar to Vogler and DeSalle (1994), Palumbi and Baker (1994) compared the overall patterns of geographic variation in nuclear intron sequences with previous mtDNA analyses and found the same pattern of geographic variability and population genetic structure among Hawaiian and Californian humpback whales, as well as similar estimates of gene flow. Performing a comparison in this study, the nuclear ITS1 region had the same pattern of high gene diversity and low nucleotide sequence diversity in Bermudian populations as did the mitochondrial control region, supporting the idea of rapid growth of Bermudian populations following founder or recolonization events.

Diversity between Bermuda and the United States

A general trend of reduced mitochondrial genetic diversity in Bermudian populations relative to the U.S. was found for all species examined. Specifically, mitochondrial control region haplotype diversities and mean nucleotide sequence diversities within Bermudian populations are lower than or equal to values estimated within the U.S. for each species with the single exception of *Holocentrus adscensionis*. However, the differences in haplotype diversity and mean nucleotide diversity between Bermuda and the U.S. were not significant.

Gene diversities and mean nucleotide sequence diversities of ITS1 within Bermudian populations do not exhibit the same general trend of reduced levels in Bermuda as was found in the control region, and are not significantly lower than those within U.S. populations. Specifically, *Haemulon sciurus*, *Holocentrus adscensionis*, and *Lagodon rhomboides* all exhibit a smaller number of alleles, lower gene diversity, and lower mean nucleotide sequence diversity in Bermuda than in the U.S., while *Haemulon aurolineatum* and *Holocentrus rufus* exhibit a greater number of alleles, higher gene diversity, and higher nucleotide sequence diversity in Bermuda than in the U.S.

If U.S. populations did indeed serve as a source for Bermudian recolonization, one would expect to see a larger reduction in both mitochondrial and nuclear diversity between Bermudian and U.S. populations provided the founder population was relatively small. A comparison of Australian bluefish (*Pomatomus saltatrix*) with populations along the Atlantic coast of the U.S. by Graves et al. (1992) provides an example of the level of reduction expected. RFLP analysis of mitochondrial DNA revealed the nucleon diversity (0.107) in Australia was five times lower than that of all other samples, and

mean nucleotide sequence diversity (0.07%) was an order of magnitude lower. In addition, the haplotypes present in Australia were unique and not found within any of the Atlantic populations. Goodbred and Graves (1996) expanded upon this study and again found that nucleon diversity in eastern Australia was very low (0.105) compared to all other populations sampled worldwide, including a population in western Australia, as was mean nucleotide sequence diversity (0.030). Eastern and western Australian bluefish were very closely related with a separation of only 0.42%. The authors suggest that the low variation present in eastern Australian populations, particularly in relation to western Australia, is consistent with the hypothesis of a recent founder event.

The lack of a significant reduction in genetic diversity in Bermudian populations does not clearly support the idea of a founder event as in the bluefish or a single recolonization event, but may in fact be indicative of a recent recolonization event followed by continued gene flow. In the neighbor-joining trees for each species using both the control region and ITS1 (Figures 8-15 and 16-23), Bermudian and U.S. samples are intermingled with one another to a large degree, with the single exception of *Haemulon aurolineatum*. A separation of the Bermudian and U.S. populations into distinct clades would be expected if the two populations had been isolated from one another. In addition, three mitochondrial control region haplotypes of *Haemulon flavolineatum* are shared between Bermuda and the U.S. while one to three nuclear ITS1 alleles are shared between Bermudian and U.S. populations of *Haemulon sciurus*, *H. flavolineatum*, *Holocentrus adscensionis*, and *Lagodon rhomboides*. There may actually be separation of the populations with no gene flow between them, but not enough time for the accumulation of sufficient variation to reveal itself in a gene tree (Benzie 1999) as

has been demonstrated in many invertebrates such as giant clams (Benzie and Williams 1997). However, it is likely that these results support the idea of continued gene flow to Bermuda from the U.S. and/or the Caribbean, which is believed to occur as rare recruitment events over evolutionary timescales (Schultz and Cowen 1994).

Critical to an evaluation of genetic diversity is a discussion of sample size. Only eight individuals of each species were sequenced from each population (seven from the U.S. population of *Lutjanus griseus* as one was removed due to misidentification). From the extremely high haplotype diversities of the mitochondrial control region, it is doubtful that all haplotypes within each population were sampled or that genetic diversity among each population was wholly represented. In fact, within eleven of the seventeen Bermudian and U.S. populations sampled, unique control region haplotypes were found in every individual. These unique haplotypes were not shared between Bermuda and the United States. In addition, one and two individuals from the U.S. populations of *Haemulon sciurus* and *H. flavolineatum*, respectively, are divergent outliers in the neighbor-joining trees using both the control region and ITS1. The minimum spanning networks reveal that the sequences of these three individuals are between 43 and 103 basepairs different from the most similar sequence found in the other individuals sampled in the U.S. In terms of ITS1 specifically, two to three clones were sequenced from each individual for most species, but only one clone was sequenced from a few individuals of *Holocentrus rufus*. All of this limits the interpretation of the results as it is clear not all the variation present in the populations has been sampled. The conclusion to be drawn is that the genetic diversity data support a recent recolonization event, but do not unambiguously support a lack of gene flow following the recolonization.

Divergence in the mitochondrial control region

If a recolonization event had occurred in Bermuda following the most recent glacial period, it would be expected that all species examined would share a similar level of divergence between U.S. and Bermudian populations correlated with approximately 18,000 years of separation. The observed values of net mean nucleotide sequence divergence for the control region in this study (0.000236 to 0.006492) vary by as much as an order of magnitude, even within the same genus. A similar level of variability in divergence values was found by Knowlton et al. (1993) between geminate pairs of *Alpheus* shrimp species across the Isthmus of Panama. Specifically, using accepted rates of cytochrome *b* evolution of 2.2 to 2.6% per million years, molecular divergence in three transisthmian sister species pairs varied nearly threefold and represented three staggered isolation events between 4.0 and 9.1 million years ago.

Correlating the estimates of divergence with time in this study can be somewhat subjective as there are several published rates of sequence evolution for the control region. Montoya-Burgos (2003) estimated an unusually slow evolutionary rate of 0.93% per million years for species of South American catfish in the genus *Hypostomus*. This estimate is based solely on a very conserved section within the center of the control region where replication begins and ends. At the opposite extreme, McMillan and Palumbi (1997) estimated a rate of mutation in a 200 bp hypervariable subsection of the 5' end of the control region near the tRNA proline gene in chaetodontids as high as 33% - 108% per million years. Using the same calibration method, they also estimated a rate of mutation of this 5' hypervariable region in cichlids between 15% and 38% per million years. However, for the purposes of this study, a rate of sequence evolution averaged

across the control region as a whole is necessary in that the entire region has been sequenced but not separated into conserved versus hypervariable sections.

In salmon and brook charr, Bernatchez and Danzmann (1993) found that the rate of mutation of the entire control region is only slightly higher than the accepted average of 2% sequence variation per million years across the mitochondrial genome as a whole. Donaldson and Wilson (1999) estimated this rate to be 3.6 ± 0.46% per million years for geminate species of snook across the Isthmus of Panama, but large variances associated with this estimate do not eliminate the possibility of lesser or greater rates. In addition, estimates for mammals (Vigilant et al. 1991), white sturgeon (Brown et al. 1993), and birds (Edwards and Wilson 1990) have fallen within a range of rates between 11.5% and 20% sequence variation per million years.

A correlation between transition/transversion ratio and the rate of control region evolution has been recently described which can aid in the choice of an appropriate rate to use for the species in this study. Bernatchez and Danzmann (1993) found a transition/transversion ratio throughout the entire control region in brook charr of 8:3 (approximately 3:1) and a similar ratio in *Poecilia reticulata*. Both fishes exhibit rates of mutation slightly higher than the average across the entire mitochondrial genome. In contrast, McMillan and Palumbi (1997) found transition/transversion ratios in an approximately 200 bp subsection of the 5' end of the control region near the tRNA proline gene as high as 25:1 in chaetodontids, species that exhibit a very high rate of control region evolution. The authors also found a transition/transversion ratio in this 5' end subsection of approximately 2:1 in fishes that exhibited rates of control region sequence evolution similar to that observed across the rest of the mitochondrial molecule.

In general, fishes with transition/transversion ratios higher than 10:1 exhibit moderate to fast control region evolution, while fishes with transition/transversion ratios closer to 3:1 exhibit mutation rates only slightly higher than the rest of the mitochondrial molecule (Bernatchez and Danzmann 1993; McMillan and Palumbi 1997).

In order to determine which rate should be used to estimate divergence time between Bermudian and U.S. populations, transition/transversion ratios of the entire control region and of an approximately 300 basepair subsection of the 5' end of the control region near the tRNA proline gene were calculated. A 300 bp section was chosen instead of the 200 bp section of McMillan and Palumbi (1997) because the specific 200 basepair subsection they used for their estimation is unclear. The majority of the variation in the hypervariable 5' end of the control region occurs within 300 bp of the tRNA proline gene (Figure 4 in Lee et al. 1995) and thus approximately 300 bp were chosen in this study in order to include the entire hypervariable region.

The three haemulid species all exhibit transition/transversion ratios near 3:1 in both the 300 bp subsection and in the control region as a whole. As Bernatchez and Danzmann (1993) describe the rate of control region evolution to be slightly higher than the accepted average of 2% per million years, the rate of 3.6 ± 0.46% estimated by Donaldson and Wilson (1999) was chosen. The remaining species (*Holocentrus adscensionis*, *H. rufus*, *Lagodon rhomboides*, and *Lutjanus griseus*) all exhibit ratios of approximately 10:1. As this is on the lower end of the range in which the rate of control region evolution is considered 'moderate to fast', divergence times were calculated using both the rate of 3.6 ± 0.46% estimated by Donaldson and Wilson (1999) for comparison with the haemulid species, and using the average accepted rate of 11.5% estimated across

several vertebrate taxa, including humans (Vigilant et al. 1991) and white sturgeon (Brown et al. 1993).

The hypothesis of a single, recent recolonization event following the end of the Wisconsin glaciation 18,000 years ago is not supported by the widely varying divergence times estimated using the mitochondrial control region. Mean divergence times appear to instead favor a hypothesis of two recolonization events. The first included *Haemulon aurolineatum*, *H. sciurus*, and *Holocentrus rufus* with mean times of divergence between 56,500 and 214,500 years ago. A second colonization appears to have occurred between 0 and 21,500 years ago including the species *Haemulon flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus*.

Correlation Between Markers

The second objective of this study was to determine if the ratio between net mean nucleotide sequence divergence estimated using the nuclear ITS1 region and that estimated using the mitochondrial control region is constant across all species examined and similar in value to ratios observed in previous studies. Montoya-Burgos (2003) found a ratio of one-third between ITS divergence values and divergence values estimated from the central conserved section of the control region in the South American catfish genus *Hypostomus*. A ratio between the rate of ITS evolution and that of the control region as a whole was described as "approximately half" by Presa et al. (2002) in a study of brown trout. Indeed, ratios of ITS divergence to that of control region divergence for each species in this study range from 0.12 to 0.99 with a mean of 0.45. Although the ratio across all species examined is not strictly constant, the mean was

similar to that observed by Presa et al. (2002), and divergence times were thus estimated for the ITS1 data using rates of control region evolution divided by two (Table 10).

Divergence Times

While not a perfect match, the divergence times estimated using ITS1 do overlap with the estimates of divergence time using the control region and the two sets of divergence times are not significantly different from one another. This similarity in estimates of divergence time using the control region and ITS1 lends support to the idea that there was not a single, recent recolonization event. Even if a different control region rate of sequence evolution is used to estimate divergence times, the ITS1 divergence estimates will remain approximately one half those of the control region. The range of divergence estimates thus support the idea that species, even within the same genus, have colonized Bermuda at different times. There does not appear to be a signal for a single, post-glacial recolonization of all the species examined following the most recent ice age.

Assuming that the appropriate rates of sequence evolution were chosen and the times of divergence are approximately correct, it would appear that some of the species did indeed begin to diverge following the last glacial period 18,000 years ago. *Haemulon flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus* have mean divergence times based on the control region of 4,000 to 21,500 years. Mean divergence times based on ITS1 for these species range from 2,000 to 20,000 years. While the range of means does extend slightly beyond the hypothesized limit of 18,000 years, the values as a whole suggest colonization of Bermuda by these four species as the last ice age ended.

It appears that the remaining species, *Haemulon aurolineatum*, *H. sciurus*, and *Holocentrus rufus* began to diverge prior to the most recent glacial period. Mean divergence times based on the control region range from 56,500 to 214,500 years while mean divergence times based on ITS1 range from 38,000 to 155,000 years. It is known that the most recent glacial period known as the Wisconsin glaciation began approximately 80,000 years before present (BP) with subsequent buildup of ice sheets and a decrease in sea level (Kennett 1982). It is unlikely that successful recruitment of colonists would have occurred during this glacial period of low sea levels and temperatures. Eleven glacial episodes have occurred within the last one million years (Kennett 1982), and thus it seems likely that colonization of these species would have occurred during one of the interglacial periods prior to the Wisconsin glaciation. A major interglacial, warming period occurred 120,000 to 130,000 years ago at which time glaciers and sea ice receded, water temperatures increased and sea level rose a peak 6 m above current levels (Kennett 1982; Alley et al. 1996; Smith-Vaniz et al. 1999). Two minor interglacial high stands in sea level also occurred at 82,000 and at 102,000 years before present (Kennett 1982). The estimated times of divergence for *Haemulon aurolineatum*, *Haemulon sciurus*, and *Holocentrus rufus* would suggest colonization during one of these interglacial, warming periods. It is most likely that recolonization occurred during the major warming period around 120,000 years ago due to its length and to the extent of the recession of glacial conditions worldwide.

In support of a colonization event prior to the Wisconsin glaciation, significant Fst values based on both the mitochondrial control region and ITS1 are found between Bermudian and U.S. populations of *Haemulon aurolineatum*, *Haemulon sciurus*, and

Holocentrus rufus which appear to have colonized approximately 120,000 years ago. The remaining species (*Haemulon flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus*), which appear to have diverged following the most recent glacial period 18,000 years ago, do not exhibit significant Fst values based on the mitochondrial control region. However, *Lutjanus griseus* and *Haemulon flavolineatum* do exhibit significant Fst values of 0.06154 and 0.23078, respectively, based on ITS1 data.

Grady et al. (2001) also found evidence of multiple colonization events to Bermuda in an analysis of four species of killifishes (*Fundulus*). Sequence variation in the mitochondrial cytochrome *b* gene and subsequent divergence estimates based on a calibrated mtDNA molecular clock were consistent with two colonization events. The first appears to have occurred over one million years ago which may have actually predated the terrestrial components of the islands, and the second appears to have occurred approximately 5,000 years ago. While these divergence times do not match those estimated in my study, the observation of multiple colonization events by Grady et al. (2001) lends support to the idea that multiple colonizations can and have occurred in the phylogeographic history of Bermudian fishes.

Interestingly, there may also be evidence of two separate colonization events within individual species in this study as well. The neighbor-joining tree and minimum spanning network for *Haemulon aurolineatum* based on ITS1 sequence data reveal two distinct clusters of Bermudian alleles, each with a central, shared allele surrounded by alleles that are only one or two basepairs different. One of these clusters is far removed from the other Bermudian alleles, separated into its own distinct, deeply divergent clade.

This more divergent clade taken separately has a mean divergence time from the U.S. alleles of approximately 617,000 years, while the remaining cluster and Bermudian alleles have a mean divergence time of 117,000 years.

Holocentrus rufus exhibits a similar separation of Bermudian alleles using ITS1 data. A deeply divergent clade of Bermudian alleles is separate from all other Bermudian and U.S. alleles in the neighbor-joining tree and forms its own cluster in the minimum spanning network. The separated clade has a mean divergence time from U.S. alleles of approximately 575,000 years, while the remaining Bermudian alleles have a mean divergence time of 9,000 years.

It is particularly interesting to note that both *Haemulon aurolineatum* and *Holocentrus rufus* have clusters of nuclear ITS1 alleles that would suggest a colonization event to Bermuda occurred approximately 600,000 years ago. The remaining alleles in *Haemulon aurolineatum* agree with the proposed colonization during a major interglacial period between 120,000 and 130,000 years ago while the remaining alleles of *Holocentrus rufus* appear to have colonized Bermuda following the Wisconsin glaciation.

This leads to the question of why some of the species in this study apparently survived the last glacial period, and perhaps prior glacial periods as well, while others did not. *Haemulon flavolineatum* appears to be a very recent colonizer of Bermuda while congeners *H. aurolineatum* and *H. sciurus* appear to have colonized prior to the last glacial period. All three are believed to have similar life history patterns, reaching maturity at approximately age two, with peak spawning occurring in the fall and winter through offshore migrations at dusk, and pelagic larval durations of 13 – 25 days (Thresher 1984; McFarland et al. 1985; Lindeman 1986; Garcia-Cagide 1986; Lindeman

et al. 2001). *H. flavolineatum* has a lower temperature limit between 14°C and 16°C (Hoss et al. 1986) and as congeners, it would be expected that *H. aurolineatum* and *H. sciurus* would share a similar range of thermal tolerance.

Holocentrus adscensionis appears to have colonized following the most recent glacial period, while *H. rufus* appears to have colonized following the previous glacial period 130,000 years ago. It may also be true that a colonization of *H. rufus* occurred around 600,000 years ago while a second colonization occurred recently. Both species share similar life history patterns, even as far as sharing a "meeki" stage of postlarval development not found in other holocentrids that may extend the pelagic larval period prior to settlement (Tyler et al. 1993). While the lower thermal minimum of *Holocentrus* has not been published to my knowledge, *H. adscensionis* has been observed as far north as inshore New York waters along with *H. vexillarius*, which would suggest that lower thermal limits may be similar to that of the haemulids (Schaefer and Doheny 1970).

Both *Lagodon rhomboides* and *Lutjanus griseus* appear to have colonized sometime during or following the most recent glacial period. These species spawn in offshore aggregations, *L. griseus* doing so on the full moon, and use estuaries as nursery grounds (Thresher 1984; Muncy 1984; Bortone and Williams 1986). *L. rhomboides* can tolerate temperatures as low as 10°C (Muncy 1984) with a critical thermal minimum of 3.4°C (Bennett and Judd 1992), while the lethal limit for *L. griseus* appears to be between 11°C and 14°C (Bortone and Williams 1986).

It is possible that *Haemulon flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus* have not been present in Bermuda as long as *Haemulon aurolineatum*, *H. sciurus*, and *Holocentrus rufus*. The former may have

colonized Bermuda for the first time as the last glacial period ended and conditions became more favorable for survival. It also seems quite likely, in that all species represented in this study appear to have similar chances for survival, that survival of *Haemulon aurolineatum*, *H. sciurus*, and *Holocentrus rufus* through the last glacial period may have been due to chance while *Haemulon flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus* were extirpated.

Anchoa

The third objective of this study was a preliminary phylogenetic investigation of three species of western central Atlantic anchovies to examine the potential site of origin of the colonists.

The closest relationship between the *Anchoa* species examined appears to be between *A. mitchilli* and a single individual of *A. hepsetus* caught in Jamaica. However, the identity of this putative *A. hepsetus* is in question. The final step used to identify the fish was a difficult choice between *A. hepsetus* and *A. colonensis*, and it has been noted that the two species are frequently confused for one another having only recently been separated. As the Jamaican sample represents only a single individual, it will not be discussed further.

The neighbor-joining trees based on both control region and ITS1 sequences (Figures 15 and 23) reveal that the endemic Bermudian *A. choerostoma* is more closely related to *A. mitchilli*, whose range is restricted to the Atlantic and Gulf coasts of the United States, than it is to *A. hepsetus*, whose range extends throughout the Caribbean. The bootstrap support for each of these clades is 100 in both the control region tree and

the ITS1 tree. While no anchovies restricted in range to the Caribbean were analyzed, the endemic Bermudian anchovy is more closely related to a species found only in the United States versus a species found throughout the Caribbean. This would suggest that colonists did not originate in the Caribbean.

Grady et al. (2001) discuss similar results among Bermudian killifishes of the genus *Fundulus*. The Bermudian endemics *F. bermudae* and *F. relictus* have a close phylogenetic relationship with the eastern North American coastal species *F. heteroclitus* and are believed to have originated through colonization from these North American populations. The relationships observed among the *Anchoa* species thus similarly suggest an origin of Bermudian colonists not in the Caribbean, but perhaps along the southeastern Atlantic coast of the United States.

Levels of taxonomic divergence

The fourth objective of my thesis is to examine the amount of genetic differentiation that separates endemic populations (conspecifics), subspecies, and species in island populations from source populations. The extremely low levels of endemism found in Bermuda in relation to levels found in other isolated island locations has been described by Smith-Vaniz et al (1999). As the shorefishes of Bermuda appear to have been derived through transport in the Gulf Stream, it would be logical to assume continued gene flow from source populations would tend to make endemism rare (Collette 1962). Collette (1962) found 4% of the total Bermudian shorefish population were endemic, Briggs (1974) claimed 5%, and Smith-Vaniz et al. (1999) found only 2.9% (10 species) of the shorefish fauna to be endemic.

Six of the species in this study represent population or conspecific comparisons: *Haemulon flavolineatum*, *H. sciurus*, *Holocentrus adscensionis*, *H. rufus*, *Lagodon rhomboides*, and *Lutjanus griseus*. The endemic Bermudian anchovy, *Anchoa choerostoma*, is considered of full species taxonomic rank, and one species, *Haemulon aurolineatum*, may represent a subspecific comparison. Ginsburg (1948) described marked differences in body depth between local populations of *Bathystoma (Haemulon) aurolineatum* in three geographic regions and classified these groups into three allopatric subspecies. One of these groups consisted only of Bermudian populations which were renamed *Bathystoma (Haemulon) aurolineatum angustum*. Collette (1962) confirmed the subspecific distinction of Bermudian populations noting the thinner body depth and a greater maximum standard length (SL) than the taxa from which they were derived. However, the subspecific distinction was not validated by Smith-Vaniz et al. (2001) and *Bathystoma (Haemulon) aurolineatum angustum* was placed back in the full species *Haemulon aurolineatum*.

A similar level of morphological distinction was found between Bermudian populations of *Hyporhamphus unifasciatus* and other populations within the western Atlantic (Banford 1993). Bermudian *Hyporhamphus unifasciatus* are at the extreme of the range of means when examining meristic characters, usually the lower end, having a smaller number of gill rakers, a thinner body depth, and a smaller size (SL) (Banford and Collette 1993; Banford 1993). Banford (1993) stressed the role of ecophenotypic plasticity in the morphological differentiation observed in Bermudian populations of both *Hy. unifasciatus* and the proposed *Ha. a. angustum*, but suggested a secondary genetic

explanation. Differences observed may reflect independent phylogenetic histories for these populations (taxa) which could be clarified through genetic analysis.

To address the case of *Haemulon aurolineatum*, divergence values were estimated between Bermudian populations and populations in the southeastern United States (from which *Ha. a. angustum* may be derived). Net mean nucleotide sequence divergence values between the two proposed subspecific populations of *H. aurolineatum* are indeed greater than those observed between conspecific populations of *H. flavolineatum*, *Holocentrus adscensionis*, *Lagodon rhomboides*, and *Lutjanus griseus*. However, both *Haemulon sciurus* and *Holocentrus rufus* exhibit divergence values between conspecific Bermudian and U.S. populations greater than that of *Haemulon aurolineatum*. The divergence value between the two proposed subspecific populations of *Haemulon aurolineatum* suggests colonization to Bermuda over 100,000 years ago as discussed previously along with *H. sciurus* and *Holocentrus rufus*. This makes it difficult to determine a level of genetic differentiation that could be used to distinguish between populations and subspecies. If all of the species examined had divergence values suggesting colonization during the earlier time period (less than 18,000 years ago) except *Haemulon aurolineatum*, it would be logical to conclude the elevated divergence value was due to prior colonization and more geological time for differentiation to the subspecific level to occur. However, this is not the case. The only apparent distinction between this subspecific comparison and the conspecific populations that also began to diverge over 100,000 years ago is in the topology of the gene trees and minimum spanning networks.

The neighbor-joining trees and minimum spanning networks for *Haemulon aurolineatum* based on both the mitochondrial control region and ITS1 show that Bermudian samples are not interspersed within U.S. samples as they are in the conspecific comparisons. The minimum spanning network based on the control region shows a single cluster of all the Bermudian haplotypes between one and two mutations removed whereas the ITS1 tree shows the majority of Bermudian alleles are in three clusters with only one or two mutations between alleles. The conspecific populations that exhibit similar divergence values to that found between the subspecies of *Haemulon aurolineatum* do show Bermudian haplotypes/alleles are intermingled with U.S. haplotypes to a much larger degree and are more than one or two mutations removed. This might suggest that the founder population of *Haemulon aurolineatum* was smaller and less genetically diverse than that of *Haemulon sciurus* or *Holocentrus rufus* and thus was more susceptible to the processes of genetic drift and speciation. Alternatively, *Haemulon sciurus* and *Holocentrus rufus* may still be experiencing small levels of continued gene flow while *Haemulon aurolineatum* populations in Bermuda may not. This could result in the greater level of morphological differentiation exhibited by *Haemulon aurolineatum* without a greater level of genetic differentiation than observed by conspecifics.

At the full species level, the endemic Bermudian anchovy, *Anchoa choerostoma*, exhibits control region divergence values from *A. mitchilli* and *A. hepsetus* of 0.120 ± 0.012 and 0.160 ± 0.013, respectively. ITS1 divergence values are 0.126 ± 0.020 and 0.188 ± 0.026, respectively. Johns and Avise (1998) compared genetic distances between sister species, congeneric species, and confamilial genera in a survey of 81 genera of fish

based on cytochrome *b* sequence data. The average value of divergence separating sister species was found to be 0.056 (Johns and Avise 1998; Colborn et al. 2001). The generally accepted rate of sequence evolution of cytochrome *b* is believed to be approximately 2% per million years (Johns and Avise 1998). *Anchoa* in this study exhibited a transition/transversion ratio between 2:1 and 4:1 which would indicate a rate of control region sequence evolution of 3.6 ± 0.46% per million years, or approximately double (1.8) that of cytochrome *b*. Using that rationale, one would expect the average value of divergence separating sister species based on the control region to be approximately 0.101 or 10%.

A. choerostoma exhibits a divergence value from its closest relative in this study of 0.120 or 12% using the control region which is similar to the estimated average value of 10% that theoretically separates sister species. Assuming that *A. mitchilli* is truly a sister species to the endemic Bermudian *A. choerostoma*, then it would appear the average level of divergence between the endemic taxonomic level of species would indeed be approximately 5% based on cytochrome *b* as described by Johns and Avise (1998), and around 10% based on the control region. However, this conclusion is tenuous as a full phylogenetic analysis has not been completed of the western central Atlantic anchovies which would clarify the true sister species of the Bermudian endemic.

Future Research

The results of this study provide a limited perspective of the phylogeography of Bermudian shorefishes. Future research should focus on increased taxonomic sampling to see if the signal for two (or more) recolonization events is supported. In addition,

increasing the number of individuals sampled, the number of clones sequenced per individual, and the number of genes analysed would increase the power associated with the study. As Smith-Vaniz et al. (1999) have noted that Bermuda shares more species in common with the Bahamas than with the southeastern Atlantic, sampling within the Caribbean would serve to answer the question of where the major source of colonists originated. Increasing the numbers of comparisons between Bermudian endemics and closely related species throughout the western central Atlantic would aid in answering this question.

Conclusions

It can be concluded that more than one interglacial recolonization of tropical and subtropical marine shorefishes to Bermuda has occurred throughout geological history. Specifically, the genetic evidence of this study suggests that there have been two, possibly three, relatively recent colonization events. The most recent occurred approximately 18,000 years ago following the end of the Wisconsin glaciation. Prior to that, colonization occurred 120,000-130,000 years ago during a major, protracted interglacial warming period. Divergence estimates to support these two colonization events were correlated between the mitochondrial control region and the nuclear ITS1 region. There is also evidence in the nuclear ITS1 data to suggest waves of colonization of both *Haemulon aurolineatum* and *Holocentrus rufus* may have occurred approximately 600,000 years ago.

A phylogenetic analysis of *Anchoa* species revealed a closer relationship between the Bermudian endemic, *A. choerostoma*, and the eastern United States species *A. mitchilli*, than between *A. choerostoma* and *A. hepsetus* which is found throughout the United States and the Caribbean. This is consistent with the hypothesis that Bermudian colonists did not originate in the Caribbean, but may have originated in populations along the southeastern Atlantic coast of the United States.

References

Alley, R; P Mayewski; D Peel; B Stauffer. 1996. Twin ice cores from Greenland reveal history of climate change, more. Earth in Space 9(2): 12-13.

Arnheim, N. 1983. Concerted Evolution of Multigene Families. In: Evolution of Genes and Proteins. M Nei and R Kohn, eds. Sinauer Associates, Inc.: Sunderland, MA. Pp 38-61.

Avise, JC. 1994. Molecular Markers, Natural History and Evolution. Chapman and Hall: New York. 511 pp.

--------. 2000. Phylogeography: The History and Formation of Species. Harvard University Press: Cambridge. 447 pp.

Baduge Suneetha, K. 2000. Intraspecific and interspecific genetic variation in selected mesopelagic fishes with emphasis on microgeographic variation and species characterization. Dr. Scient. Dissertation. Department of Fisheries and Marine Biology, University of Bergen (Norway). 34 pp.

Bennett, WA; FW Judd. 1992. Factors affecting the low-temperature tolerance of Texas pinfish. Transactions of the American Fisheries Society 121(5): 659-666.

Benzie, JAH. 1999. Genetic structure of coral reef organisms: ghosts of dispersal past. American Zoologist 39:131-145.

Benzie, JAH; ST Williams. 1997. Gene flow among giant clam (Tridacna maxima) populations in the west Pacific is not consistent with dispersal by present-day ocean currents. Evolution 51:768-783.

Bermingham, E; HA Lessios. 1993. Rate variation of protein and mitochondrial DNA evolution as revealed by sea urchins separated by the Isthmus of Panama. Proceedings of the National Academy of Sciences, USA 90: 2734-2738.

Bermingham, E; SS McCafferty; AP Martin. 1997. Fish Biogeography and Molecular Clocks: Perspectives from the Panamanian Isthmus. In: Molecular Systematics of Fishes. TD Kocher, CA Stepien, eds. Academic Press: San Diego. Pp 113-128.

Bernatchez, L; RG Danzmann. 1993. Congruence in control-region sequence and restriction-site variation in mitochondrial DNA of brook charr (*Salvelinus fontinalis* Mitchill). Molecular Biology and Evolution. 10(5): 1002-1014.

Bortone, SA; JL Williams. 1986. Species profiles: Life histories and environmental requirements of coastal fishes and invertebrates (South Florida): Gray, Lane, Mutton, and Yellowtail Snappers. U.S. Fish and Wildlife Service. 30 pp.

Bowen, BW; JC Avise. 1990. Genetic structure of Atlantic and Gulf of Mexico populations of sea bass, menhaden, and sturgeon: influence of zoogeographic factors and life-history patterns. Marine Biology 107: 371-381.

Bowen, BW; AL Bass; LA Rocha; WS Grant; DR Robertson. 2001. Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global scale. Evolution 55(5): 1029-1039.

Brogan, MW. 1994. Distribution and retention of larval fishes near reefs in the Gulf of California. Marine Ecology Progress Series 115: 1-13.

Brown, DD; PC Wensink; E Jordan. 1972. A comparison of the ribosomal DNAs of *Xenopus laevis* and *Xenopus mulleri*: the evolution of tandem genes. Journal of Molecular Biology 63: 57-73.

Brown, J; A Colling; D Park; J Phillips; D Rothery; J Wright. 1998. Ocean Circulation. G Bearman, ed. Open University, Walton Hall, Milton Keynes MK7 6AA, and Butterworth-Heinemann: Oxford. 238 pp.

Brown, JR; AT Beckenbach; MJ Smith. 1993. Intra-specific DNA sequence variation of the mitochondrial control-region of the white sturgeon (*Acipenser transmontanus*). Molecular Biology and Evolution 10: 326-241.

Brown, WM. 1983. Evolution of Animal Mitochondrial DNA. Evolution of Genes and Proteins. M Nei and R Kohn, eds. Sinauer Associates, Inc.: Sunderland, MA. Pp. 62-88.

Buckler, ES IV; TP Holtsford. 1996. *Zea* ribosomal repeat evolution and substitution patterns. Molecular Biology and Evolution 13(4): 623-632.

Colborn, J; RE Crabtree; JB Shaklee; E Pfeiler; BW Bowen. 2001. The evolutionary enigma of bonefishes (*Albula* spp.): cryptic species and

ancient separations in a globally distributed shorefish. Evolution 55(4): 807-820.

Cowen, RK; CB Paris; KMM Lwiza; DB Olson. 2000. Long distance dispersal versus local retention as a means of replenishing Caribbean marine fish populations. International Council for the Exploration of the Sea Copenhagen (Denmark) Theme Session on Spatial and Temporal Patterns in Recruitment Processes. 6 pp.

Cronin, MA; WJ Spear; RL Wilmot; JC Patton; JW Bickham. 1993. Mitochondrial DNA variation in chinook (*Oncorhynchus tshawytscha*) and chum salmon (*O. keta*) detected by restriction enzyme analysis of polymerase chain reaction (PCR) products. Canadian Journal of Fisheries and Aquatic Sciences 50: 708-715.

Donaldson, KA; RR Wilson, Jr. 1999. Amphi-Panamic geminates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. Molecular Phylogenetics and Evolution 13(1): 208-213.

Dowling, TE; C Moritz; JD Palmer; LH Rieseberg. 1996. Nucleic acids III: analysis of fragments and restriction sites. In: Molecular Systematics. DM Hillis, C Moritz, BK Mable, eds. Sinauer Associates, Inc.: Sunderland, MA. Pp. 249-320.

Edwards, SV; AC Wilson. 1990. Phylogenetically informative length polymorphism and sequence variability in mitochondrial DNA of Australian songbirds (*Pomatostomus*). Genetics 126: 695-711.

Elder, JF Jr.; BJ Turner. 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. The Quarterly Review of Biology 70(3): 297-320.

Epifanio, CE; RW Garvine. 2001. Larval transport on the Atlantic continental shelf of North America; a review. Estuarine, Coastal and Shelf Science 52: 51-77.

Garcia-Cagide, A. 1986. Reproductive characteristics of the bluestriped grunt, *Haemulon sciurus*, in the eastern zone of the Batabano Gulf, Cuba. Reporte de Investigacion del Instituto de Oceanologia, Academia de Ciencias de Cuba. Havana. No.48. 28 pp.

Goggin, CL. 1994. Variation in the two internal transcribed spacers and 5.8S ribosomal RNA from five isolates of the marine parasite *Perkinsus* (Protista, Apicomplexa). Molecular and Biochemical Parasitology 65(1): 179-182.

Goodbred, CO; JE Graves. 1996. Genetic relationships among geographically isolated populations of bluefish (Pomatomus saltatrix). Marine and Freshwater Research 47: 347-355.

Grady, JM; DK Coykendall; BB Collette; JM Quattro. 2001. Taxonomic diversity, origin, and conservation status of Bermuda killifishes (*Fundulus*) based on mitochondrial cytochrome b phylogenies. Conservation Genetics 2: 41-52.

Graves, JE; JR McDowell; AM Beardsley; DR Scoles. 1992. Stock structure of the bluefish *Pomatomus saltatrix* along the mid-Atlantic coast. Fishery Bulletin 90: 703-710.

Harris, DJ; KA Crandall. 2000. Intragenomic variation within ITS1 and ITS2 of freshwater crayfishes (Decapoda: Cambaridae): Implications for phylogenetic and microsatellite studies. Molecular Biology and Evolution 17(2): 284-291.

Hedgecock, D. 1986. Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? Bulletin of Marine Science 39(2): 550-564.

Hillis, DM; BK Mable; C Moritz. 1996. Applications of Molecular Systematics. In: Molecular Systematics. DM Hillis, C Moritz, BK Mable, eds. Sinauer Associates, Inc.: Sunderland, MA. Pp 515- 543.

Hillis, DM; MT Dixon. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. The Quarterly Review of Biology 66(4): 411-453.

Hoss, DE; DS Peters; WF Hettler. 1986. Effects of lowered water temperature on the survival and behavior of juvenile French grunt, *Haemulon flavolineatum*. Bulletin of Marine Science 39(1): 134-138.

Johannes, RE. 1978. Reproductive strategies of coastal marine fishes in the tropics. Environmental Biology of Fishes 3(1): 65-84.

Jones, GP; MJ Millcich; MJ Eemslie; C Lunow. 1999. Self-recruitment in a coral reef fish population. Nature 402: 802-804.

Kaufman, J; S Olson; J Pangrossi. 1994. "MacVector," v.7.0 Oxford Molecular Group plc, 2000. Symantec Corporation.

Keffer, T; DG Martinson; BH Corliss. 1988. The position of the Gulf Stream during Quarternary glaciations. Science 241: 440-442.

Kennett, JP. 1982. Marine Geology. Prentice-Hall, Inc: Englewood Cliffs, NJ. 813 pp.

Knowlton, N; LA Weigt, LA Solorzano; DK Mills; E Bermingham. 1993. Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. Science 260: 1629-1632.

Knowlton, N; LA Weigt. 1998. New dates and new rates for divergence across the Isthmus of Panama. Proceedings of the Royal Society of London, Series B: Biological Sciences 265: 2257-2263.

Kumar, S; K Tamura; IB Jakobsen; M Nei. 2001. MEGA2.1: Molecular Evolutionary Genetics Analysis software. Arizona State University, Tempe, Arizona, USA.

Lee, W-J; J Conroy; WH Howell; TD Kocher. 1995. Structure and evolution of teleost mitochondrial control regions. Journal of Molecular Evolution 41: 54-66.

Leis, JM. 1991. The pelagic stage of reef fishes: the larval biology of coral reef fishes. In: The Ecology of Fishes on Coral Reefs. PF Sale, ed. Academic Press, Inc.: San Diego. Pp 183-230.

Lessios, HA. 1979. Use of Panamanian sea urchins to test the molecular clock. Nature 280: 599-601.

Lessios, HA; BD Kessing; DR Robertson. 1998. Massive gene flow across the world's most potent marine biogeographic barrier. Proceedings of the Royal Society of London, Series B: Biological Sciences 265: 583-588.

Li, W-H. 1993. So, what about the molecular clock hypothesis? Current Opinion in Genetics and Development 3: 896-901.

Lindeman, KC. 1986. Development of larvae of the french grunt, *Haemulon flavolineatum*, and comparative development of twelve species of western atlantic Haemulon (Percoidei, Haemulidae). Bulletin of Marine Science 39(3): 673-716.

Lindeman, KC; TN Lee; WD Wilson; R Claro; JS Ault. 2001. Transport of larvae originating in southwest Cuba and the Dry Tortugas: Evidence for partial retention in grunts and snappers. Proceedings of the Gulf and Caribbean Fisheries Institute 52: 732-747.

Martin, AP; GJP Naylor; SR Palumbi. 1992. Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature 357: 153-155.

Mayer, MS; PS Soltis. 1999. Intraspecific phylogeny analysis using ITS sequences: insights from studies of *Streptanthus glandulosus* complex (Cruciferae). Systematic Botany 24: 47-61.

McBride, RS; KW Able. 1998. Ecology and fate of butterflyfishes, *Chaetodon* spp., in the temperate, western North Atlantic. Bulletin of Marine Science 63(2): 401-416.

McFarland, WN; EB Brothers; JC Ogden; MJ Shulman; EL Bermingham. 1985. Recruitment patterns in young French grunts, *Haemulon flavolineatum* (Family Haemulidae) at St. Croix. U.S. Virgin Islands Fisheries Bulletin 83: 413-426.

McMillan, WO; SR Palumbi. 1997. Rapid rate of control-region evolution in Pacific butterflyfishes (Chaetodontidae). Journal of Molecular Evolution 45: 472-484.

Milliman, JD; KO Emery. 1968. Sea levels during the past 35,000 years. Science 162: 1121-1123.

Montoya-Burgos, JI. 2003. Historical biogeography of the catfish genus *Hypostomus* (Siluriformes; Loricariidae), with implications on the diversification of Neotropical ichthyofauna. Molecular Ecology 12: 1855-1867.

Muncy, RJ. 1984. Species profiles: Life histories and environmental requirements of coastal fishes and invertebrates (Gulf of Mexico): Pinfish. U.S. Fish and Wildlife Service. 30 pp.

Muss, A; DR Robertson; CA Stepien; P Wirtz; BW Bowen. 2001. Phylogeography of *Ophioblennius*: the role of ocean currents and geography in reef fish evolution. Evolution 55(3): 561-572.

Palumbi, SR. 1994. Genetic divergence, reproductive isolation, and marine speciation. Annual Review of Ecology and Systematics 25: 547-572.

--------. 1996. What can molecular genetics contribute to marine biogeography? An urchin's tale. Journal of Experimental Marine Biology and Ecology 203: 75-92.

Palumbi, SR; CS Baker. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. Molecular Biology and Evolution 11(3): 426-435.

Powers, TO; TC Todd; AM Burnell; PCB Murray; CC Fleming; AL Szalanski; BA Adams; TS Harris. 1997. The rDNA internal transcribed spacer region as a taxonomic marker for nematodes. Journal of Nematology 29(4): 441-450.

Presa, P; BG Pardo; P Martinez; L Bernatchez. 2002. Phylogeographic congruence between mtDNA and rDNA ITS markers in brown trout. Molecular Biology and Evolution 19(12): 2161-2175.

Riffel, M; A Schreiber. 1995. Coarse-grained population structure in Central European sculpin (*Cottus gobio* L.): Secondary contact or ongoing genetic drift? Journal of Zoological Systematics and Evolutionary Research 33(4): 173-184.

Rozas, J; R Rozas. 1999. DnaSP version 3.1: An integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15(2): 174-175.

Sambrook, J; DW Russell. 2001. Molecular Cloning: A Laboratory Manual. 3rd Edition. Cold Spring Harbor Laboratory Press: New York.

Schaefer, RH; T Doheny. 1970. A first record of the squirrelfish from New York waters. New York Fish and Game Journal 17(1): 60-61.

Scheltema, RS. 1968. Dispersal of larvae by equatorial ocean currents and its importance to the zoogeography of shoal-water tropical species. Nature 217: 1159-1162.

--------. 1995. The relevance of passive dispersal for the biogeography of Caribbean mollusks. American Malacological Bulletin 11(2): 99-115.

Schizas, NV; GT Street; BC Coull; GT Chandler; JM Quattro. 1999. Molecular population structure of the marine benthic copepod *Microarthridiion littorale* along the southeastern and Gulf coasts of the USA. Marine Biology 135: 399-405.

Schneider, S; D Roessli; L Excoffier. 2000. Arlequin ver 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schultz, ET; RK Cowen. 1994. Recruitment of coral-reef fishes to Bermuda: local retention or long-distance transport? Marine Ecology Progress Series. 109: 15-28.

Shulman, MJ; E Bermingham. 1995. Early life histories, ocean currents, and the population genetics of Caribbean reef fishes. Evolution 49(5): 897-910.

Smith-Vaniz, WF; BB Collette; BE Luckhurst. 1999. Fishes of Bermuda: History, Zoogeography, Annotated Checklist, and Identification Keys. American Society of Ichthyologists and Herpetologists Special Publication No. 4.

Swearer, SE; JE Caselle; DW Lea; RR Warner. 1999. Larval retention and recruitment in an island population of coral-reef fish. Nature 402: 799-802.

Taberlet, P. 1996. The use of mitochondrial DNA control region sequencing in conservation genetics. In: Molecular Genetic Approaches in Conservation. Oxford University Press: New York. Pp 125-142.

Thompson, JD; DG Higgins; TJ Gibson. 1994. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and matrix choice. Nucleic Acids Research 22: 4673-4680.

Thresher, RE. 1984. Reproduction in Reef Fishes. T. F. H. Publications, Inc. Ltd.: Neptune City, NJ. 399 pp.

Turner, TF; JC Trexler; DN Kuhn; HW Robison. 1996. Life-history variation and comparative phylogeography of darters (Pisces: Percidae) from the North American central highlands. Evolution 50(5): 2023-2036.

Tyler, JC; GD Johnson; EB Brothers; DM Tyler; CL Smith. 1993. Comparative early life histories of western Atlantic squirrelfishes (Holocentridae): Age and settlement of rhynchichthys, meeki, and juvenile stages. Bulletin of Marine Science 53(3): 1126-1150.

Vawter, AT; R Rosenblatt; GC Gorman. 1980. Genetic divergence among fishes of the eastern Pacific and the Caribbean: support for the molecular clock. Evolution 34(4): 705-711.

Vigilant, LM; M Stoneking; H Harpending; K Hawkes; AC Wilson. 1991. African populations and the evolution of human mitochondrial DNA. Science 236: 1503-1507.

Vogler, AP; R DeSalle. 1994. Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle *Cincindela dorsalis*. Molecular Biology and Evolution 11(3): 393-405.

Vita

Kelly R. Johnson

Born in Fargo, North Dakota, on November 18, 1977. Graduated from Fargo South High School, Fargo, North Dakota, in 1996. Graduated with honors from North Dakota State University in 2000 with a Bachelor of Science in Zoology, concentration in Fisheries and Wildlife Management. Entered the Master of Science program at the Virginia Institute of Marine Science/School of Marine Science of the College of William and Mary in Virginia in 2000.